Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4 1-187-07-4406-10

Anomalous Mental Phenomena: Selected Papers

Compiled By:

The Cognitive Sciences Laboratory

24 June 1991

Science Applications International Corporation

An Employee-Owned Company

5150 El Camino Real, Suite B-31, Los Altos, California 94022 (415) 960-5910

Other SAIC Offices: Albuquerque, Boston, Colorado Springs, Dayton, Huntsville, Las Vegas, Los Angeles, McLean, Oak Ridge, Orlando, Palo Alto, San Diego, Seattle, Tucson

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

This document is made available through the declassification efforts and research of John Greenewald, Jr., creator of:

The Black Vault

The Black Vault is the largest online Freedom of Information Act (FOIA) document clearinghouse in the world. The research efforts here are responsible for the declassification of hundreds of thousands of pages released by the U.S. Government & Military.

Discover the Truth at: **http://www.theblackvault.com**

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

I INTRODUCTION

In this volume, we present a selected set of papers on, and/or in support of, anomalous mental phenomena. No section could possibly be complete; however, we have chosen papers that are representative of their particular sections. The sections, which are separated by blue sheets, are as follows:

	Section	Number of Papers
I	Introduction	
II	Meta-analyses of Anomalous Mental Phenomena	8
III	Main-stream Publications	7
IV	Anomalous-mental-phenomena Journal Publications	6
V	Magnetoencephalography	3
VI	Physics	6

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

II META-ANALYSES OF ANOMALOUS MENTAL PHENOMENA

As in all behavioral sciences, replication of experiments in anomalous mental phenomena (AMP) is critical before any putative effects can be verified as part of nature. Because of the complex nature of most behavioral experiments, drawing conclusions from a body of similar experiments has been problematical. Meta-analysis, however, is a relatively new statistical approach that has been specifically designed to address the particular difficulties inherent in the behavioral sciences.

The papers in this section have been selected because they represent all such analyses of a substantial portion of the published AMP literature to date. Through replication and meta-analysis, the general scientific community will have tools with which to judge the claims of the AMP literature.

The number that appears in the upper right–hand corner of the first page for each publication is keyed to the following descriptions:

1. Utts, J., "Successful Replication Versus Statistical Significance," *Journal of Parapsychology,* Vol. 52, pp. 305-320, (December, 1988). By defining, in statistical terms, the meaning of replication for few–σ effects, Utts, a Professor of Statistics from the University of California at Davis, sets the statistical basis for meta-analysis.
2. Honorton, C., "Error Some Place!" *Journal of Communication*, pp. 103-116, (Winter, 1975). This paper predates the development of formal meta-analysis, but Honorton provides a critical review of all the ESP card-guessing experiments from 1934 to 1939. The paper includes a description of the claims and counter-claims surrounding the controversy of the day.
3. Honorton, C. and Ferrari, D. C., "'Future telling:' A meta-Analysis of Forced-Choise Precognition Experiments, 1935-1987," *Journal of Parapsychology*, Vol. 53, pp. 282-308, (December, 1989). Using the full complement of meta-analytical tools, Honorton provides a critical review of all the ESP experiments during which the target material (i.e., usually ESP cards) is generated *after* the guess has been recorded.
4. Honorton, C., Berger, R. E., Varvoglis, M. P., Quant, M., Derr, P., Schechter, E., I., and Ferrari, D. C., "Psi Communication in the Ganzfeld," *Journal of Parapsychology*, Vol. 54, pp. 99-137, (June, 1990). This paper provides a meta-analysis of Ganzfeld experiments (i.e., a form of anomalous cognition). The database is comprised of 11 series for a total of 355 individual trials.
5. Radin, D. I. and Nelson, R. D., "Evidence for Consciousness-Related Anomalies in Random Physical Systems," *Foundations of Physics*, Vol. 19, No. 12, pp. 1499-1514, (December, 1989). Radin and Nelson analyze over 800 experiments that claim evidence for mental human-machine interactions (i.e., anomalous perturbation). After a careful analysis, which includes accounting for experiment flaws, they conclude that there is substantial statistical evidence to support the claim.
6. Honorton, C., Ferrari, D. C., and Bem, D. J., "Extraversion and ESP Performance: Meta-Analysis and a New Confirmation," *Proceedings of the Parapsychological Association 33rd Annual Convention,* Chevy Chase, MD, (August, 1990). In an important link to traditional psychological experimentation, this paper provides a meta-analysis for the correlation of ESP performance and a traditional personality variable, extraversion.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

7. Rosenthal, R., "Meta-Analytic Procedures and the Nature of Replication: The Ganzfeld Debate," *Journal of Parapsychology*, Vol. 50, pp. 319-336, (December, 1986). Rosenthal, a professor of psychology at Harvard University, is one of the early developers of the meta-analysis techniques. In this paper, he comments about the Ganzfeld controversy.

8. Utts, J., "Replication and Meta-Analysis in Parapsychology," Accepted for publication in *Statistical Sciences*. In this paper, Utts, provides an independent and objective overview of the AMP meta-analyses that follow.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Journal of Parapsychology, Vol. 52, December 1988

SUCCESSFUL REPLICATION VERSUS STATISTICAL SIGNIFICANCE

BY JESSICA UTTS

ABSTRACT: The aim of this paper is to show that successful replication in parapsychology should not be equated with the achievement of statistical significance, whether at the .05 or at any other level. The p value from a hypothesis test is closely related to the size of the sample used for the test; so a definition of successful replication based on a specific p value favors studies done with large samples. Many "nonsignificant" studies may simply be ones for which the sample size was not large enough to detect the small magnitude effect that was operating. Conversely, "significant" studies may result from a small but conceptually insignificant bias, magnified by a very large sample.

The paper traces the history of the definition of statistical significance in parapsychology and then outlines the problems with using hypothesis-testing results to define successful replications, especially when applied in a cookbook fashion. Finally, suggestions are given for alternative approaches to looking at experimental data. These include calculating statistical power before doing an experiment, using estimation instead of, or in conjunction with, hypothesis testing, and implementing some of the ideas from Bayesian statistics.

Replication is a major issue in parapsychology. Arguments about whether a given research paradigm has been successful tend to focus on what the replication rate has been. For example, the recent review of parapsychology by the National Research Council includes statements such as "...of these 188 [RNG] experiments with some claim to scientific status, 58 reported statistically significant results (compared with the 9 or 10 experiments that would be expected by chance)" (Druckman & Swets, 1988, p. 185). In each section, the report critically evaluates "significant" experiments and ignores "nonsignificant" experiments. The extent to which nonsignificant experiments are ignored is exemplified by the following oversight, in which the total number of studies is equated with the number of "successful" studies: "Of the *thirteen* scientifically reported experiments [of remote viewing], *nine* are classified as successful in their outcomes by Hansen et al. ... As it turns out, all but one of the *nine* scientifically reported studies of remote viewing suffer from the flaw of sensory cueing" (p. 183, emphasis added). Apparently the authors decided that the four experiments that did not attain a p value of .05 or less did not even warrant acknowledgment.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The practice of defining a successful replication as an experiment that attains a p value of .05 or less is common in parapsychology, psychology, and some other disciplines that use statistics. However, like many other conventions in science, it is based on a series of historical events rather than on rational thought. In this paper, I will trace some of the history leading to this definition of a "successful" experiment, outline some problems with this approach, and suggest some methods that parapsychologists should consider in addition to the usual hypothesis-testing regimen. Rao (1984) and Honorton (1984) have discussed similar problems and solutions in the context of psi experiments.

History

It has not always been the case among parapsychologists that an experiment was deemed successful if it reached a significance level of $p = .05$. In 1917, John Edgar Coover, who was the Thomas Welton Stanford Psychical Research Fellow at Stanford University from 1912 to 1937, published a book with the results from several experiments he had conducted up to that time (Coover, 1917/1975). Although hypothesis testing as we know it today had not yet been formalized, he essentially conducted tests on many facets of this data and found no evidence for psi that was convincing to him. His conclusions regarding these results are typified by an example he gave in which the hit rate for 518 trials was 30.1%, when 25% was expected by chance (exact p value = .00476):

> We get 0.9938 [p-value = 1 − 0.9938 = 0.0062] for the probability that chance deviations will not exceed this limit [of 30.1 percent]. . . . Since this value, then, lies within the field of chance deviation, although the probability of its occurrence by chance is fairly low, it cannot be accepted as a decisive indication of some cause beyond chance which operated in favor of success in guessing. (p. 82)

He then revealed what level of evidence would convince him that nonchance factors were operating: ". . . if we meet the requirement of a degree of accuracy usual in scientific work by making P = 0.9999779, when absolute certainty is P = 1, then [there is] satisfactory evidence for some cause in addition to chance" (p. 83). In other words, he was defining significance with a p value of 2.21×10^{-5}.

Coover was not alone in requiring that results conform to arbitrarily stringent significance levels. In 1940, when Rhine et al. published *Extra-Sensory Perception After Sixty Years*, they included the following definitions in the glossary:

> p-value = probability of success in each trial
>
> SIGNIFICANCE: When the probability that chance factors alone produced a given deviation is sufficiently small to provide relative certainty that chance is not a reasonable expectation, the deviation is *significantly* above or below the chance level. Among ESP results, this is arbitrarily taken to mean a deviation in the expected direction such that the critical ratio is 2.5 times the standard deviation (or four times the probable error) or greater. (p. 423–424)

Thus, significance was defined by $z \geq 2.5$, or $p \leq .0062$.

Seventeen years later, in their book *Parapsychology: Frontier Science of the Mind,* Rhine and Pratt (1957) suggested that .01 was the appropriate threshold:

> In order for such judgments to have the necessary objectivity, a *criterion of significance* is established by practice and general agreement among the research workers in a particular field. . . . Most workers in parapsychology accept a probability of .01 as the criterion of significance. (p. 186)

Finally, the *Journal of Parapsychology* has included a definition of *significance* in its glossary for many years, but the appropriate p-value has fluctuated back and forth between .01 and .02, finally settling at .02 in 1968. The following are excerpts from those glossaries:

December 1949: "A numerical result is significant when it equals or surpasses some criterion of degree of chance improbability. Common criteria are: a probability value of .01 or less."

March 1950 to June 1957: "The criterion commonly used in this Journal is a probability value of .02 or less."

September 1957: "The criterion commonly used in this Journal is P = .01."

December 1957 to December 1967: "The criterion commonly used in parapsychology today is a probability value of .01 or less."

March 1968 to December 1986: "The criterion commonly used in parapsychology today is a probability value of .02 (odds of 50 to 1 against chance) or less. . . . Odds of 20 to 1 (probability of .05) are regarded as strongly suggestive."

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

March 1987: The term *significance* no longer appears in the glossary.

By the mid-1980's, despite the value of .02 given in the *Journal of Parapsychology*, significance seemed to have been determined to correspond to a p value of .05. For example, in their bibliography of remote-viewing research, Hansen, Schlitz, and Tart (1984) claim: "We have found that more than half (fifteen out of twenty-eight) of the published formal experiments have been successful, where only one in twenty would be expected by chance." As mentioned in my introduction, .05 was the value used by the National Research Council in their recent evaluation of parapsychology. Both Hyman (1985) and Honorton (1985) used .05 as the criterion for a successful ganzfeld study. In discussing the Schmidt REG experiments, Palmer (1985) implicitly used .05 as the cut-off for significance by observing: "Based on Z-tests . . . 25 of the 33 (76%) were significant at the .05 level, two-tailed. In two of the seven non-significant studies. . . ." (p. 102).

This definition of significance is obviously not unique to parapsychology. A popular introductory textbook in psychology states that:

> Psychologists used a statistical inference procedure that gives them an estimate of the probability that an observed difference could have occurred by chance. This computation is based on the size of the difference and the spread of the scores. By common agreement, they accept a difference as "real" when the probability that it might be due to chance is less than 5 in 100 (indicated by the notation $p < .05$). A **significant difference** is one that meets this criterion. . . . With a statistically significant difference, a researcher can draw a conclusion about the behavior that was under investigation. (Zimbardo, 1988, p. 54)

Given the weight that has been attached to .05 as the criterion for significance, one would think that it resulted from careful consideration of the issue by statisticians and psychologists. Unfortunately, such is not the case. Its roots apparently lie in the following passage published in 1926 by one of the founders of modern statistics, Sir Ronald A. Fisher:

> It is convenient to draw the line at about the level at which we can say: "Either there is something in the treatment, or a coincidence has occurred such as does not occur more than once in twenty trials." . . . If one in twenty does not seem high enough odds, we may, if we prefer it, draw the line at one in fifty (the 2 per cent point), or one in a hundred (the 1 per cent point). Personally, the writer prefers to set a low standard of significance at the 5 per cent point, and ignore entirely all results which fail to reach that level. A scientific fact should be regarded as experimentally established only if a properly designed experiment *rarely fails* to give this level of significance. (Fisher, 1926, p. 504; also quoted in Savage, 1976, p. 471)

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Thus began the belief that an experiment is successful only if the null hypothesis can be rejected using $\alpha = 0.05$. As an immediate consequence of this belief, Fisher and his followers created tables of F statistics that included values only for tail areas of .05 and .01. Since researchers did not have access to computer algorithms to determine intermediate p values, success came to be measured in terms of these two values alone.

Problems with Hypothesis Testing

Misconceptions about p Values

Most modern research reports include p values instead of simply discussing whether an experimental result is significant at a prespecified level. Although this is somewhat better than the old method of "one star or two" (corresponding to a significant result at .05 or .01, respectively), it is still a misleading way to examine experimental results.

The problem is that many researchers interpret p values as being related to *the probability that the null hypothesis is true.* Even some sophisticated researchers tend to think that an extremely small p value must correspond to a very large effect in the population and that a large p value (say $> .10$) means that there is no effect. In other words, *the size of the p value is incorrectly interpreted as the size of the effect.* It should be interpreted as the probability of observing results as extreme or more so than those observed, *if* there is no effect.

To see how arbitrary it is to base a decision about the truth or falsity of a statement on a p value, consider a binomial study based on a sample of size n which results in $z = 0.30$, p value $= .38$, one-tailed. One would probably abandon the hypothesis under study and decide not to pursue the given line of research. Now suppose that the study had been run with a sample of size $100n$ instead and resulted in the exact same proportion of hits. Then we would find $z = 3.00$, p value $= .0013$. These results would be regarded as highly significant!

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

As another example, consider a chi square test for randomness based on a sequence of n numbers, each of which can take the values 1, 2, ... 10. Suppose that the test results in a chi-square value of 11.0, $df = 9$, p value = 0.28. Now suppose the sequence was three times as long but the proportions of each digit remained the same. Then each term in the numerator of the chi-square statistic would be multiplied by 3^2, whereas each term in the denominator would only be multiplied by 3. The degrees of freedom would not change, but the new result would be $\chi^2 = 33.0$, $df = 9$, p value = .00013. In the first case, the conclusion would be that the sequence was sufficiently random, yet a sequence three times as long with the same pattern would be seen to deviate considerably from randomness!

This problem was recognized more than 50 years ago by Berkson (1938):

> We may assume that it is practically certain that any series of real observations does not actually follow a normal curve *with absolute exactitude* in all respects, and no matter how small the discrepancy between the normal curve and the true curve of observations, the chi-square P will be small if the sample has a sufficiently large number of observations in it.
>
> If this be so, then we have something here that is apt to trouble the conscience of a reflective statistician using the chi-square test. For I suppose it would be agreed by statisticians that a large sample is always better than a small sample. If, then, we know in advance the P that will result from an application of a chi-square test to a large sample, there would seem to be no use in doing it on a smaller one, but since the result of the former test is known, it is no test at all. (pp. 526–527, emphasis in original)

Replication

Very often researchers simply do not understand the connection between the p value and the size of the sample. For example, Rosenthal and Gaito (1963) asked nine faculty members and ten graduate students in a university psychology department to rate their degree of belief or confidence in results of hypothetical studies with various p values and with sample sizes of 10 and 100. Given the same p value, one should have more confidence in a study with a *smaller* sample because it would take a larger underlying effect to obtain the small p value for a small sample. Unfortunately, these

respondents demonstrated that they were far more likely to believe results based on the large sample when the p values were the same. (For a discussion of this example and some other problems with hypothesis testing in psychology, see Bakan, 1967.)

One consequence of this misunderstanding is that researchers misinterpret what constitutes a "successful replication" of an experiment. Tversky and Kahneman (1982) asked 84 members of the American Psychological Association or the Mathematical Psychology Group the following question:

> Suppose you have run an experiment on 20 subjects, and have obtained a significant result which confirms your theory ($z = 2.23$, $p < .05$, two-tailed). You now have cause to run an additional group of 10 subjects. What do you think the probability is that the results will be significant, by a one-tailed test, separately for this group? (p. 23)

The median answer given was .85. Only 9 of the 84 respondents gave an answer between .40 and .60. Assuming that the value obtained in the first test was close to the true population value, the probability of achieving a p value $\leq .05$ on the second test is actually only about .47. This is because the sample size in the second study is so small. The effect would have to be quite large in order to be detected with such a small sample.

In the same survey, Tversky and Kahneman also asked:

> An investigator has reported a result that you consider implausible. He ran 15 subjects, and reported a significant value, $t = 2.46$. Another investigator has attempted to duplicate his procedure, and he obtained a nonsignificant value of t with the same number of subjects. The direction was the same in both sets of data. You are reviewing the literature. What is the highest value of t in the second set of data that you would describe as a failure to replicate? (p. 28)

The majority of respondents considered $t = 1.70$ as a failure to replicate. But if the results from both studies are combined, then (assuming equal variances) the result is $t = 2.94$, $df = 29$, p value = .003. The paradox is that the new study *decreases* faith in the original result if viewed separately but *increases* it when combined with the original data!

This misunderstanding about replication is quite prevalent in the psi literature, as demonstrated by the emphasis on successful replication, where success is defined in terms of a specific p value, regardless of sample size. As an example of how unnecessarily dis-[...]ers, I have shown elsewhere (Utts,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

1986) that if the true hit rate in a binomial study (such as a ganzfeld experiment) is actually 33%, and 25% is expected by chance, then a study based on a sample of size 26 should be expected to be "successful" ($p \leq .05$) only about one fifth of the time. Even a study based on a sample of size 100 should be "successful" only about half of the time. It is no wonder that there are so many "unsuccessful" attempts at replication in psi.

As another example of the paradoxical nature of this definition of replication, consider the "unsuccessful" direct-hit ganzfeld studies covered by the meta-analyses of Hyman (1985) and Honorton (1985). Using those studies with $p(\text{hit}) = .25$, there were 13 out of 24 that were nonsignificant, $\alpha = 0.05$, one-tailed. (See Honorton, p. 84, Table A1.) But when these 13 "failures" are combined, the result is 106 hits out of 367 trials, $z = 1.66$, $p = .0485$!

Problems with Point Null Hypotheses

A point null hypothesis is one that specifies a particular value ("point") as the one being tested. Most hypothesis testing is done with point null hypotheses. The problem with this approach is that any given hypothesis is bound to be false, even if just by a minuscule amount. For example, in a coin-tossing experiment, the null hypothesis is that the coin is fair, that is to say, H_o: $P = .5000000$. This is never precisely true in nature. All coins and coin-tossers introduce a slight bias into the experiment. This slight bias can produce a very small p value if the sample size is large enough. If, for example, the true probability of heads is .5001, and the observed proportion of heads falls right at this value, then the null hypothesis will be rejected at .05 if the sample size is at least 6.7×10^7. As long as there is any bias *at all*, the p value can be made arbitrarily small by taking a large enough sample.

In practice, this problem was rarely serious before it became possible to collect large amounts of data rapidly using computers. Statisticians have often used ESP as an example of one of the few cases where it really is possible to specify an exact value for the null hypothesis. But even this view is changing, as shown by this comment from a recent issue of a popular statistics journal:

> It is rare, and perhaps impossible, to have a null hypothesis that can be exactly modeled as $\theta = \theta_o$. One might feel that hypotheses such as
>
> H_o: A subject has no ESP, or
>
> H_o: Talking to plants has no effect on their growth,
>
> are representable as exact (and believable) point nulls, but, even here, minor biases in the experiments will usually prevent exact representations as points. (Berger & Delampady, 1987, p. 320)

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

In summary, hypothesis testing as it is currently formulated tends to be a misleading approach to examining data. Small samples tend to lead to "nonsignificant" studies, whereas large samples can lead to extremely small p values, even if the null hypothesis is only slightly wrong. Many researchers do not understand the meaning of a p value and do not understand how closely replication issues are tied to sample size. Arguments about replication should not be based on p values alone.

Solutions

Power Calculations

If a hypothesis test is to be done at all, a researcher should at least determine in advance whether it is likely to be successful. The statistical power of a test is the probability that the null hypothesis will be rejected. It obviously depends on what the true underlying state of nature is. Because this information cannot be known (or there would be no point in doing the experiment), it is a good idea to look at power for a variety of possibilities *before* conducting the experiment. The results will tell you whether you are likely to be able to reject the null hypothesis, using the sample size you have planned, for specific values of the magnitude of the effect.

Statistical power is a function of the sample size, the true underlying magnitude of the effect, the level of significance for which the experiment would be considered a success, and the method of analysis used. It does not depend on the data.

As an example, suppose you are planning to conduct a test of the hypothesis H_o: $P = .25$ using a series of 10 independent trials. Power calculations would proceed as follows:

1. Find the cutoff point for the number of hits that would lead to rejection of H_o. In this case, the p value for 5 hits is .08, and for 6 hits it is .02, so 6 hits would probably be required to reject the null hypothesis.

2. Power for a specific alternative is the probability that the null hypothesis would be rejected if that alternative value is true. In this case, power = P(6 or more hits). This can be computed directly, using the binomial formula, for any specified hit rate. Here are some examples:

Hit rate	Power = P(6 or more hits)
0.30	.047
0.33	.073
0.40	.166
0.50	.377

Notice that even if the true hit rate is 50% instead of the chance level of 25%, the chances of a "successful" replication are poor, that is, only 37.7%. In most psi applications, 30% or 33% is probably a more realistic approximation to the true hit rate, so there would be a very small chance of having this experiment succeed with only 10 trials.

As a second example, suppose you are planning to run the same experiment with 100 trials and are planning to use the normal approximation instead of an exact test. Further, suppose you will reject the null hypothesis if $z \geq 1.645$, where z is the usual critical ratio, corrected for continuity: $z = (\text{number of hits} - 0.5 - 25) / \sqrt{(100 \times .25 \times .75)} = .23(\text{number of hits} - 25.5)$. Using simple algebra, note that $z \geq 1.645$ when the number of hits $\geq$ 32.65. Thus, the null hypothesis will be rejected if there are 33 or more hits, so power = P(33 or more hits). Computing this for the same hypothetical hit rates as in the previous example gives:

Hit rate	Power = P(33 or more hits)
0.30	.289
0.33	.538
0.40	.939
0.50	.9998

Now there is a more reasonable chance for a successful study, although it is still only 29% even if the true hit rate is 30%.

For studies in which the null hypothesis does not involve a single value, it can be more difficult to compute power because it is not so easy to specify a reasonable alternative. In these cases, it is still possible to look at the p value that can be expected if psychic functioning were to occur at specified levels for the sample size planned. For example, McClenon and Hyman (1987) conducted a remote-viewing study with eight trials, one for each of eight subjects, and used the preferential-ranking method of Solfvin, Kelly, and Burdick (1978) on the subject rankings. Each subject was asked to rank-order eight choices of potential targets as compared to the response he or she had produced. By chance, the average rank should be 4.5. If psychic functioning had reduced the average rank to 4.0, the p value would have been .298, not significant. Even if the average rank had been reduced to 3.5, the study would still not have been significant, p value = .126. The average rank would have to be 3.0 before this study would achieve a significant result. A parapsychologist experienced in remote viewing should be able to determine in advance whether such a study would be likely to be successful with such a small sample.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The lesson here is that a "nonsignificant" study may be nothing more than a study with low power. Before investing time and money in a new study, it should be determined whether it is likely to succeed if psychic functioning is operating at a given level.

Estimation

An approach that avoids many of the problems with hypothesis testing is to construct a "confidence interval" or an "interval estimate" for the magnitude of an effect. This is done by computing an interval of values that almost certainly covers the true population value. The degree of certainty is called the *confidence coefficient* and is specified by the researcher. Common values are 95% and 99%.

As an example, consider a binomial study with 100 trials that results in 35 hits. Using the normal approximation, one would expect the proportion of hits in the sample to be within 1.96 standard deviations of the true hit rate 95% of the time. The appropriate standard deviation for the proportion P of hits is $\sqrt{P(1 - P)/n}$. Thus, a 95% confidence interval for the true hit rate is found by adding and subtracting 1.96 of these standard deviations to the proportion of hits observed in the sample. The resulting interval in this case is 0.35 − 0.09 to 0.35 + 0.09, or 0.26 to 0.44. This tells us that with a fair amount of certainty (95%), the true hit rate is covered by the interval from 0.26 to 0.44. For the same proportion of hits in a study with 1,000 trials, the interval would be from 0.32 to 0.38. The larger the sample size, the shorter the width of the interval.

Consider two studies designed to test H_0: $P = .5$:

	Study 1	Study 2
z	3.60	2.40
p value	.0004	.0164
n	1,000	100

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Which study provides more convincing evidence that there is a strong effect? In keeping with the results of Rosenthal and Gaito (1963) discussed earlier, most people would say that the first study shows a stronger effect, both because the p value is smaller and because it is based on a larger sample. In fact, the opposite is true. The number of hits for the two studies are 557 (55.7%) and 62 (62%), respectively; the smaller study had a higher hit rate. The 95% confidence intervals for the hit rates in the two studies are (0.53 to 0.59) and (0.53 to 0.72), respectively, so in both studies we are relatively sure that the hit rate is at least 53%, but in the second study it could be as high as 72% whereas in the first it is probably no higher than 59%.

In studies with huge sample sizes, confidence intervals make it evident that an infinitesimal p value does not correspond to an effect of large magnitude. For example, consider a study based on 100,000 trials and designed to test H_0: $P = .50$. Suppose there were 50,500 hits. Then $z = 3.16$, and the p value is 7.9×10^{-4}. But what does this mean in practical terms? A 95% confidence interval for the true hit rate is from 0.5019 to 0.5081. Thus, it appears that the true hit rate is indeed different from 0.50, but reporting the results in this way makes it clear that the magnitude of the difference is very small. The reader can decide whether an effect of this size has any meaning in the context of the experiment.

In summary, confidence intervals are preferable to hypothesis tests for the following reasons:

1. They show the *magnitude* of the effect.
2. They show that the accuracy of the conclusion is highly dependent on the sample size.
3. They remove the focus from decision making, which is arbitrary at best because of sample size problems.
4. They highlight the distinction between *statistical* significance and *practical* significance.
5. They allow the reader of a research report to come to his or her own conclusion.

Meta-Analyses

Meta-analytic techniques may be viewed by some parapsychologists as the solution to studying the issue of replication. Even though these techniques can address the replication issue in useful ways, they also contain some dangerous pitfalls. For example, both Hyman (1985) and Honorton (1985) used "vote-counting" in their meta-analyses of the ganzfeld data base. In other words, they tallied the number of significant studies in the data base. This procedure inherits all of the problems associated with the original determination of whether a study was "significant" in the first place. A series of studies, each with low power, may all be determined to be nonsignificant, when the combined data may lead to an extremely significant result. Conversely, a series of studies based on large samples may all be significant, but the magnitude of the effect may be very small. A vote-count showing that most studies are significant could mislead researchers into believing that there was a large effect.

The concept of effect size was introduced to account for the fact that individual study results are highly dependent on sample size. Estimating the effect sizes for a series of studies and seeing whether they are similar is a useful way of studying replication. However, examining only the effect size for an individual study does not give any indication of the accuracy of the result. This should be done in conjunction with some estimate of the accuracy of the result, such as a confidence interval.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Bayesian Methods

Many statisticians believe that the conceptual framework of hypothesis testing and interval estimation is philosophically incorrect. Rather, they start by assigning prior probabilities, based on subjective belief, to various hypotheses, and then combine these "priors" with the data to compute final or "posterior" probabilities for the hypotheses. This is called the Bayesian approach to statistics. An introduction to the ideas of Bayesian analysis can be found in Berger and Berry (1988) or Edwards, Lindman, and Savage (1963). A more technical reference is Berger (1985).

Berger and Berry (1988), in a recent article in *American Scientist*, discussed the use of Bayesian methods instead of classical methods:

> The first step of this demonstration is to calculate the actual probability that the hypothesis is true in light of the data. This is the domain of Bayesian statistics, which processes data to produce "final probabilities" ... for hypotheses. Thus, the conclusion of a Bayesian analysis might be that the final probability of H is 0.30.
>
> The direct simplicity of such a statement compared with the convoluted reasoning necessary to interpret a P-value is in itself a potent ar-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

> gument for Bayesian methods. Nothing is free, however, and the elegantly simple Bayesian conclusion requires additional input. To obtain the final probability of a hypothesis in light of the experimental data, it is necessary to specify the probability of the hypothesis before or apart from the experimental data.
>
> Where does this initial probability come from? The answer is simple. It must be subjectively chosen by the person interpreting the data. A person who doubts the hypothesis initially might choose a probability of 0.1; by contrast, someone who believes in it might choose 0.9. (p. 162)

They then provide an example of testing the hypothesis H: $P = .5$, where P is the proportion of hits expected in a binomial experiment. Suppose that in 17 trials there are 13 successes (76.5%). Then the p value is .049, two-tailed. Unless, of course, the experiment was designed to stop at the fourth failure instead of at the 17th trial. Then the p value, with the identical data, would only be .021. Such problems arise with classical methods, but not with Bayesian methods.

Using the Bayesian approach, suppose that one's prior belief that H is true is 50%. If H isn't true, the prior belief is that the true value of P is equally likely to be anywhere between $0.5 - c$ and $0.5 + c$ (where c is some constant), but could not possibly be farther than that from 0.5. The choice of c represents prior opinion about the strength of the effect, if there is one. Choosing $c = 0.1$ (the effect isn't likely to be very strong even if it exists) results in a final probability of 0.41 for H (given that there were 13 successes in 17 trials), whereas choosing $c = 0.4$ results in a probability of 0.21 for H. In other words, the final degree of belief in H is dependent on one's prior belief about the strength of the effect. It also depends on prior opinion about the veracity of H, and on the observed data.

One reason that Bayesian methods are not more widely used is that they are often difficult to apply. Another reason is that researchers are uncomfortable with having to specify subjective degrees of belief in their hypotheses. This approach makes particular sense for parapsychology, however, because most researchers have strong opinions about the probability that psi is real, and these opinions play a central role in how psi researchers and critics evaluate the evidence. Posterior probabilities in Bayesian analyses are a function of both the prior probabilities and the strength of the evidence; it may be informative to formalize these opinions and to see how much evidence would be needed to increase the posterior probability of a psi hypothesis to a non-negligible level when the prior probability was close to zero.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

References

Bakan, D. (1967). *On method: Toward a reconstruction of psychological investigation.* San Francisco: Jossey-Bass, Inc.

Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis.* New York: Springer-Verlag.

Berger, J. O., & Berry, D. A. (1988, March–April). Statistical analysis and the illusion of objectivity. *American Scientist,* pp. 159–165.

Berger, J. O., & Delampady, M. (1987). Testing precise hypotheses. *Statistical Science,* 2(3), 317–334.

Berkson, J. (1938). Some difficulties of interpretation encountered in the application of the chi-square test. *Journal of the American Statistical Association,* 33, 526–542.

Coover, J. E. (1975). *Experiments in psychicial research.* New York: Arno Press. (Originally published 1917)

Druckman, D., & Swets, J. A. (1988). *Enhancing human performance.* Washington, D. C.: National Academy Press.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological Review,* **70**, 193–242.

Fisher, R. A. (1926). The arrangement of field experiments. *Journal of the Ministry of Agriculture of Great Britain,* **33**, 503–513.

Hansen, G. P., Schlitz, M. J., & Tart, C. T. (1984). Bibliography, remote-viewing research, 1973–1981. In R. Targ & K. Harary, *The mind race* (pp. 265–269). New York: Villard Books.

Honorton, C. (1984). How to evaluate and improve the replicability of parapsychological effects. In B. Shapin & L. Coly (Eds.), *The repeatability problem in parapsychology* (pp. 238–255). New York: Parapsychology Foundation, Inc.

Honorton, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. *Journal of Parapsychology,* **49**, 51–91.

Hyman, R. (1985). The ganzfeld psi experiment: A critical appraisal. *Journal of Parapsychology,* **49**, 3–49.

McClenon, J., & Hyman R. (1987). A remote viewing experiment conducted by a skeptic and a believer. *Zetetic Scholar,* Nos. 12/13, 21–33.

Palmer, J. (1985). An evaluative report on the current status of parapsychology. U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, VA.

Rao, K. R. (1984). Replication in conventional and controversial sciences. In B. Shapin & L. Coly (Eds.), *The repeatability problem in parapsychology,* (pp. 22–41). New York: Parapsychology Foundation, Inc.

Rhine, J. B., & Pratt, J. G. (1957). *Parapsychology: Frontier science of the mind.* Springfield IL: Charles C. Thomas.

Rhine, J. B., Pratt, J. G., Stuart, C. E., Smith, B. M., & Greenwood, J. A. (1940). *Extra-sensory perception after sixty years.* Boston: Bruce Humphries.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Rosenthal, R., & Gaito, J. (1963). The interpretation of levels of significance by psychological researchers. *Journal of Psychology*, 55, 33–38.

Savage, L. J. (1976). On rereading R. A. Fisher. *Annals of Statistics*, 4, 441–500.

Solfvin, G. F., Kelly, E. F., & Burdick, D. S. (1978). Some new methods for preferential-ranking data. *Journal of the American Society for Psychical Research*, 72, 93–110.

Tversky, A., & Kahneman, D. (1982). Belief in the law of small numbers. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Utts, J. M. (1986). The ganzfeld debate: A statistician's perspective. *Journal of Parapsychology*, 50, 393–402.

Zimbardo, P. G. (1988). *Psychology and life*. Glenview, IL: Scott, Foresman and Co.

Division of Statistics
University of California
Davis, CA 95616

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Paranormal Communication

"Error Some Place!"

by Charles Honorton

Review of the ESP controversy traces debate from statistical and methodological issues to the a priori *critique and the paradigm of "normal science."*

Asked his opinion of ESP, a skeptical psychologist once retorted, "Error Some Place!" I believe he was right, but for the wrong reasons. Western science has always been ambivalent toward the mental side of reality, and it is perhaps not surprising that the occurrence of "psychic" phenomena is one of the most controversial topics in the history of science.

The first serious effort toward scientific examination of psi claims was undertaken by the Society for Psychical Research (SPR), founded in London in 1882 for the purpose of "making an organized and systematic attempt to investigate the large group of phenomena designated by such terms as mesmeric, psychical, and spiritualistic." The SPR leadership included many distinguished scholars of the period, and similar organizations quickly spread to other countries, including the American Society for Psychical Research, founded in New York in 1885 under the aegis of William James, who himself took an active role in early investigations of mediumistic communications.

These turn-of-the-century investigators focused much of their attention on authenticating individual cases of spontaneous experiences suggestive of psi communication. While a great deal of provocative material was carefully examined and reported (e.g., 13), the limitations inherent in the case study approach prohibited definitive conclusions. However thoroughly authenticated, spontaneous cases cannot provide adequate assessment of such potential sources of contamination as chance coincidence, unconscious inference and sensory leakage, retroactive falsification, or deliberate fraud.

Charles Honorton is director of research in the Division of Parapsychology and Psychophysics, Department of Psychiatry, Maimonides Medical Center, Brooklyn, N.Y.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Early experimental approaches primarily involved the "telepathic" reproduction of drawings at a distance (62). While often striking correspondences were obtained, the experimental conditions did not usually provide for random selection of target (stimulus) material, and were not always totally adequate with respect to the possibility of sensory leakage, intentional or otherwise.

Neither the spontaneous case studies nor the early experimental efforts made much impact upon the scientific community, though they drew critical comment from prominent period scientists. "Neither the testimony of all the Fellows of the Royal Society, nor even the evidence of my own senses," proclaimed Helmholtz, "would lead me to believe in the transmission of thought from one person to another independently of the recognized channels of sense." Thomas Huxley declined an invitation to participate in some of the early SPR investigations, saying he would sooner listen to the idle gossip of old women.

The rudiments of an experimental methodology for testing psi were suggested three centuries ago by Francis Bacon.

In *Sylva Sylvarum,* a work published posthumously, Bacon discussed "experiments in consort, monitory, touching transmission of spirits and forces of imagination." He suggested that "the motions of shuffling cards, or casting of dice" could be used to test the "binding of thoughts. . . . The experiment of binding of thoughts should be diversified and tried to the full; and you are to note whether it hit for the most part though not always" (2).

The application of probability theory to the assessment of deviations from theoretically expected chance outcomes was introduced to psychical research in 1884 by the French Nobel laureate, Charles Richet, in experiments involving card-guessing. The popularity of card-guessing as an experimental methodology was greatly influenced by the work of J. B. Rhine and his associates at Duke University in the early 1930s. Rhine (50) devised a standard set of procedures around a simplified card deck containing randomized sequences of five geometric forms (circle, cross, wavy lines, square, and circle). These "ESP cards" were prepared in packs of 25, and each "run" through the pack was associated with a constant binomial probability of 1/5, since subjects were not given trial-by-trial feedback. Providing the experimental conditions were adequate to eliminate illicit sensory cues, recording errors, and rational inference, statistically significant departures from binomial chance expectation were interpreted as indicating extrasensory communication.

Initially, "telepathy" tests consisted of having a subject in one room attempt to identify the order of the cards as they were observed by an "agent" in another room. In "clairvoyance" tests, the subject attempted to "guess" the order of the cards directly, as they lay concealed in an opaque

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

container or in another room, without an agent. "Precognition" tests, introduced somewhat later (59), required the subject to make anticipatory guesses of the card order before the pack was shuffled or otherwise randomized.

Rhine introduced the term "ESP" in his first major report on the Duke University work in 1934 (50). He reported a total of 85,724 card-guessing trials, carried out with a wide variety of subjects and under a wide range of test conditions. The results as a whole were astronomically significant, though informal exploratory trials were indiscriminately pooled with those carried out under more carefully controlled conditions. The best-controlled work during this period was the Pearce-Pratt distance series of clairvoyance tests (58), in which the subject, Pearce, located in one building, attempted to identify the order of the cards as they were handled, but not viewed, by Pratt, the experimenter, located in another building. The level of accuracy obtained in this series of 1,850 trials was associated with a probability of 10^{-22}.

As a stimulant to experimental research, Rhine's work had unprecedented influence. For the first time a common methodology was adopted and employed on a large scale by a number of independent and widely separated investigators. For the first time, also, the scientific community was confronted with a body of data, collected through conventional methods, which it could no longer ignore—nor too hastily accept. The wide-scale adoption of the card-guessing methodology was accompanied by a plethora of critical articles, challenging almost every aspect of the evaluative techniques and the experimental conditions. During the period between 1934 and 1940, approximately 60 critical articles by 40 authors appeared, primarily in the psychological literature. While card-guessing is no longer the primary methodology in experimental parapsychology, the questions which arose over its use are of equal relevance to the more sophisticated approaches used today.

The first major issue concerned the validity of the assumption that the probability of success in the card-guessing experiments was actually 1/5.

If chance expectation is other than 1/5, the significance of the observed deviations would obviously be in doubt. This issue was quickly resolved by mathematical proof and through empirical "cross-checks," a form of control series in which responses (guesses) were deliberately compared with target orders for which they were not intended (e.g., responses on run n_1 matched with the target sequence for run n_2). Empirical cross-checks were reported for 24 separate experimental series involving a total of 12,228 runs (305,700 individual trials). While the actual experimental run scores (e.g., guesses on run n_1 compared to targets for run n_1) were highly significant and yielded a mean scoring rate of 7.23/25, the control cross-check

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

scores were in all cases nonsignificant, with a mean scoring rate of 5.04 (43).

Several critics questioned the applicability of the binomial distribution as a basis for assessing the statistical significance of ESP card-guessing data. Willoughby (78) proposed the use of an empirical control series, but later withdrew the suggestion after comparing the two methods (79). Alternative methods of deriving the probable error and recommendations for using the empirical standard deviation were also proposed and later withdrawn (21, 22). Concern over this issue diminished and was generally abandoned following the publication of a large chance control series involving half a million trials and demonstrating close approximation to the binomial model (12).

Another question arose about whether the binomial model provides sufficient approximation to the normal distribution to allow use of normal probability integral tables for determination of significance levels (17). Stuart and Greenwood (73) showed that when the normal distribution is used as an approximation to the binomial model, discrepancies are important only with cases of borderline significance and few trials.

The use of the binomial critical ratio (z) to evaluate the significance of the ESP card-guessing deviations was generally approved by professional statisticians (6, 20). Fisher (10), however, commented that high levels of statistical significance should not be accepted as substitutes for independent replication. In another vein, Huntington (20) asked, "If mathematics has successfully disposed of the hypothesis of chance, what has psychology to say about the hypothesis of ESP?"

The most frequently expressed methodological concern was the possibility of some form of "sensory leakage," giving the ESP subject enough information about the targets to account for significant, extrachance results

As early as 1895, two Danish psychologists, Hansen and Lehmann (16), reported that with the aid of parabolic reflectors subjects could detect digits and other material silently concentrated upon by an agent. In these experiments, the subject and agent sat with their heads close to the foci of two concave mirrors. While the agent concentrated on the number, he made a special effort to keep his lips closed. Under these conditions, the subjects were frequently successful in identifying the number. These results were interpreted by Hansen and Lehmann as supporting the hypothesis of "involuntary whispering." The utilization of subtle sensory cues was demonstrated in a careful investigation by S. G. Soal of a stage "telepathist" (66). There were also reports, such as the case of "Ilga K.," a mentally retarded Latvian child who could read any text, even in a foreign language when someone stood behind her, reading "silently." Experiments with dictaphone recordings revealed that "Ilga" was responding to very slight auditory cues (3).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

able to the ESP hypothesis made 71.5 percent more errors of commission (increasing ESP scores), while those who were unfavorable to the ESP hypothesis made 100 percent more errors of omission (decreasing ESP scores). Murphy (37) reported an analysis of 175,000 trials from experiments reporting positive evidence for ESP and found only 175 errors (0.10 percent). Greenwood (12) reported only 90 recording errors in rechecking his 500,000-trial control study, of which 76 were errors of omission.

Some critics also alleged that improper selection of data could account for experimental successes. This could be done in several ways: (a) selection of subjects; (b) selection of particular blocks of data out of larger samples; (c) selection of one of several forms of analysis; and (d) selective reporting of particular studies. The questions raised have sometimes been stated cynically in the form, "Parapsychologists must run 100 subjects before they find one with 'ESP'." As if in defense against this charge, a number of the reported studies specifically stated that all of the data collected were included in the analysis (see 43, pp. 118–124, Table 12).

Concerning selection of subjects, Warner (76) suggested two criteria: first, results of "poor" subjects must be included up to the point when they are discontinued since it does not matter how many trials a given *subject* makes as long as all of the *trials* (for all subjects) are included; second, exclude *all* preliminary trials (for both "good" and "poor" subjects) and use preliminary screening studies to select "good" candidates for formal work. These criteria were generally endorsed by the chief critics of the period (e.g., 23).

The question of *post hoc* selection of analyses was not a point of serious concern in the period between 1934 and 1940, though it is relevant to the assessment of some of the process-oriented investigations reported more recently. The question of whether nonsignificant studies were withheld from publication involves an issue which is of great concern to the behavioral sciences as a whole (70, 81) and one which is difficult to accurately assess since there is no way of knowing how many studies may have been withheld from publication because their results failed to disconfirm the null hypothesis.

Several studies of American Psychological Association publication policies (4, 70, 81) indicate that experimental studies in general are more likely to be published if the null hypothesis is rejected at the conventional .05 and .01 alpha levels than if it is not rejected. These studies also indicate that a negligible proportion of published studies are replications. Bozarth and Roberts (4), in a survey of 1,334 articles from psychological journals, found that 94 percent of the articles involving statistical tests of significance reported rejection of specific null hypotheses; only eight articles (less than 1 percent) involved replications of previously published studies.

With respect to the implications of such selection for the ESP hypothesis, there are two partial answers. First, considering the degree of critical interest which prevailed in the 1930s, it seems unlikely that nonsignificant findings would have been repressed during this period; second, the high levels

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

It is clear that at least some of the early exploratory series reported in Rhine's monograph were open to criticism for inadequate controls against sensory cues. While Rhine did not base major conclusions on such poorly controlled data, inclusion of them in his monograph provided a ready target for critical reviewers and sidetracked discussion away from the better controlled work, such as the Pearce-Pratt series, which was not susceptible to explanation by sensory cues.

Defects in an early commercial printing of ESP cards were reported by several investigators (18, 25). It was found that the cards were warped and could under certain conditions be identified from the back. This discovery circulated widely for a time as an explanation of all successful (i.e., statistically significant) experimental series. The parapsychologists retorted that defective cards had not been employed in any of the experiments reported in the literature and that, in any case, they could not account for results from studies involving adequate screening with such devices as opaque envelopes, screens, distance, or work involving the precognition paradigm in which the target sequences were not generated until after the subject had made his responses (53, 54, 72).

By 1940 nearly one million experimental trials had been reported under conditions which precluded sensory leakage. These included five studies in which the target cards were enclosed in opaque sealed envelopes (41, 45, 46, 54, 59), 16 studies employing opaque screens (7, 8, 11, 19, 33, 34, 35, 38, 41, 42, 44, 45, 46, 59, 71), ten studies involving separation of subjects and targets in different buildings (50, 51, 52, 53, 34, 32, 8, 77, 61, 60), and two studies involving precognition tasks (59, 75). These data are summarized in Table 1. The results were independently significant in 27 of the 33 experiments. By the end of the 1930s there was general agreement that the better-controlled ESP experiments could not be accounted for on the basis of sensory leakage.

The hypothesis that significant "extrachance" deviations in ESP experiments might be attributable to motivated scoring errors was investigated in several studies. In one investigation (26), 28 observers recorded 11,125 mock ESP trials. Of these, 126 (1.13 percent) were misrecorded. Observers favor-

Table 1: ESP card-guessing experiments (1934-1939) excluding sensory cues[a]

Method	Studies	N (Trials)	Mean/25	p<	z
"Clairvoyance" paradigm, stimuli in sealed, opaque envelopes	5	129,775	5.21	4.0×10^{-11}	7.56
"Clairvoyance" paradigm, stimuli concealed by opaque screens	16	497,450	5.44	2.0×10^{-11}	31.03
Distance[b]	10	164,475	5.37	10^{-11}	15.00
Precognition paradigm[c]	2	115,330	5.15	2.95×10^{-6}	5.09

[a] References given in text.
[b] Includes work with both "telepathy" and "clairvoyance" paradigms
[c] Stimuli generated after subjects made their responses

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Two recent examples, one involving cancer research (74) and the oth involving parapsychology (57), serve to remind us of the importance cross-validation in the assessment of *any* experimental finding. In both cases, it should be added, the fraudulent acts were detected in-house, the researchers themselves. The point is that in the final analysis an expe mental finding is of value and is to be taken seriously only to the extent that it leads to further inquiry. To regard any experiment as an end in itself is to remove it from the domain of experimental science. It is obvio that hypothetical construct, such as ESP, cannot be validated by any isolated experiment, no matter how well controlled it might be. Independent replication is a necessary prerequisite.

The claim that psi phenomena operate outside the framework of physical probabili has been a major source of priori *arguments against acceptance of ESP.*

It has been suggested that to accept ESP requires the rejection physics. This is absurd, and it is worth noting that such arguments have usually been advanced and defended by psychologists rather than physicists.

The debate over the incompatibility of physics and ESP has been conducted almost exclusively within the framework of nineteenth-century deterministic physics, wherein the ultimate constituent of physical reali was still believed to be solid matter. Inasmuch as modern microphysics has exorcised the material out of matter and deals with processes which on our macrophysical level of sensory perception are every bit as erratic ar anomalous as ESP, the *a priori* claim that ESP violates specifiable laws physics can no longer be considered to be of more than historical interest.

ESP and other psi phenomena, while no longer incompatible wi physics, are not yet accounted for by physics; but then, neither are t more familiar processes of memory and conscious experience. Indeed, the transformation of "raw feels" into conscious experience is no less a problem for the neurophysiologists of today than it was for the speculative philc ophers of classical antiquity. Sir John Eccles, among others, has repeatedly warned, "We should not pretend that consciousness is not a mystery."

The ESP controversy illustrates several features of the paradigmat view of science developed by Thomas Kuhn (28). Normal science, according to Kuhn, is essentially a clean-up operation, constrained by a bro theoretical framework, or paradigm, which defines the boundaries of legi mate inquiry. Paradigms are scientific world views which provide coherence and structure and determine the types of questions to be posed of natur as well as the manner in which answers are sought. Normal science is th a process of paradigm-articulation, rather than of discovery. Within the paradigm structure of normal science, observations which conflict with

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

remaining [illegible] scientist is ignorance, ignorance concerning the work [illegible] ning its implications. The typical scientist contents [illegible] ing . . . some criticism that at most applies to a small [illegible] published studies. But these findings (which challenge [illegible] of space and time) are—if valid—of enormous imp[illegible] ey ought not to be ignored.

Following [illegible] nt on miracles, Price asserted that "ESP is "incompatible [illegible] entific theory," and that it is therefore more parsimonious [illegible] parapsychologists cheat than that ESP [illegible] real phenomen[illegible] ed, "My opinion concerning the findings of the parapsych[illegible] many of them are dependent on clerical and statistical errors and unintentional use of sensory clues, *and that all extra-chance results not so explicable are dependent on deliberate fraud or mildly abnormal mental conditions*" (47, p. 360). This extraordinary critique and the ensuing discussion in *Science* (5, 36, 48, 55, 5[illegible], 65) were widely reviewed. As Meehl and Scriven (40) pointed out, Price's argument rests on two highly questionable assumptions, namely that contemporary scientific knowledge is complete, and that ESP necessarily conflicts with it. Seventeen years later, in an "Apology to Rhine and Soal," Price retracted his accusations of investigator fraud (49).

Very similar arguments have, however, been made [illegible] by the British [illegible] critic C. E. M. Hansel [illegible] began his examination of the ESP hypothesis by sugg[illegible] arguments [illegible] time and effort [illegible] experiments. [illegible] *a priori* argu[illegible] *in advance* that [illegible] cannot occur.[6]

Because of [illegible] unlikelihood" of ESP, [illegible] of the literature [illegible] primarily on the possibility of [illegible] or investigators. He [illegible] in depth four experiments [illegible] as providing the [illegible] of ESP: the already [illegible] distance series [illegible] Woodruff (44) series [illegible] and Soal's work [illegible] Stewart and Basil Shackleton (60) as [illegible] a more recent [illegible] and Bowden [illegible] Hansel [illegible] case, how [illegible] committed (by [illegible] Pratt-Woodruff [illegible] series and by the subjects [illegible] Pratt and Soal [illegible] experiments). He gave no direct evidence [illegible] fraud was committed [illegible] experiments but said, "If the result could have arisen through a [illegible] experiment must be considered unsatisfactory proof of ESP, [illegible] *is finally decided that such* [illegible] *was* [illegible] *used*" (1[illegible], p. 1[illegible]).

Hansel's [illegible] inasmuch as he quite properly insists that no single experiment [illegible] conclusive, then proceeds to show that none is, given the [illegible] possibility of fraud by subjects or investigators. Hansel's [illegible] more than 2[illegible]0 pages of careful [illegible] was that [illegible] were not "fraud-proof" and therefore not conclusive proof [illegible]

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 2: Breakdown of experimental ESP studies (1934-1939)

	N (Studies)*	N studies reported significant ($p < .01$)	% signif.
Duke group	17	15	88
Non-Duke	33	20	61
Total	50	35	70

* Includes all English-language studies involving assessment of statistical significan[ce] of data, 1934–1939 inclusive.

x^2 (Duke vs. non-Duke × significant vs. nonsignificant) = 1.70 (1 df)

sequences have been "doctored" prior to publication in order to remo[ve] certain nonrandom features; this practice, according to Spencer Brown, makes such sequences nonrandom and invalidates the use of standa[rd] significance tests; (b) the source of some random number sequences invol[ve] randomizing machines which utilize the unpredictability of human behavior when examined for microscopic variation; such variation, sa[ys] Spencer Brown, may be predictable enough to account for observed ano[m]alies in random sequences, as well as some of the significant results reported in ESP guessing experiments; (c) Spencer Brown produces evidence to show that anomalous (significant) departures from probability theory can [be] obtained by matching columns of random numbers (39).

A detailed examination of these points was undertaken by Scott (64). With respect to "doctored" sequences, Scott showed that the maximu[m] error due to rejected (edited) sequences would not affect interpretatio[n] of results which are more than marginally significant and could, in fact, increase the likelihood of making a Type II error. On the hypothesis th[at] ESP results are due to some kind of hyper-regularity affecting both t[he] target sequence and the response (guess) sequence simultaneously and similarly, Scott makes the point that this would lead to the expectation of similar results from matching *any* set of humanly produced random [se]quences. The cross-check type of control series and the Greenwood chance control series described earlier demonstrate that this is not the case. The anomalies reported by Spencer Brown (68), obtained by arbitrarily matchi[ng] columns of random numbers, have been criticized on the basis of *post h[oc]* selection (40) and illustrate not that there are fundamental defects in probability theory, but rather that significant deviations from chance c[an] occur in any data where hypotheses and analyses are not specified in advan[ce].

The most recent phase of the ESP controversy centers on the hypothesis of investigator fraud. This argument was most forcefully presented in [a] lead article in *Science*, entitled "Science and the Supernatural," by G. [R.] Price (47), who began with the following observations:

> *Believers in psychic phenomena . . . appear to have won a decis[ive] victory and virtually silenced opposition. . . . This victory is the res[ult] of an impressive amount of careful experimentation and intelligent argumentation. . . . Against all this evidence, almost the only defe[nse]*

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

of significance attached to some of the reported ESP investigations would necessitate postulating astronomical numbers of "chance" trials in order to dilute the overall deviations to chance. To take one example, consider the Pearce-Pratt series of 1,850 trials which yielded $p = 10^{-22}$. As Soal and Bateman (66) pointed out, it is difficult to believe that 10^{10} (ten thousand million) sets of 1,850 trials could have possibly been carried out between 1934 and 1940 (or, for that matter, since 1940). But, as Soal and Bateman suggest, ". . . if we posit this absurd estimate as an upper limit [with overall chance totals], that would still give us odds of 10^{10} . . . against the supposition that the Pearce-Pratt results were a run of pure luck."

The possibility of obtaining significant "extrachance" results by stopping an experimental series at "favorable" points was also raised (9, 31). While this "optional stopping" hypothesis was generally agreed to be of significance only in cases of marginally significant results, it led to the adoption of several procedural modifications: specification of the total number of trials in advance of data collection, or accumulation of data in blocks of predetermined size.

The possibility was raised by several critics that hand-shuffled cards may display a tendency to stick together or otherwise produce patterns which could produce spurious results (24, 82). While the cross-check type of control series, described earlier, failed to reveal any evidence of patterning, there was a general trend away from hand shuffling in the later published studies, which utilized tables of prepared random numbers as a basis for generating target sequences.

There was—and is (e.g., 15)—a rather widespread belief that most of the evidence supporting the ESP hypothesis originated in the Duke University studies and that most independent replications by other investigators were nonconfirmatory. A survey of the published literature between 1934 and 1940 fails to support this claim. Table 2 shows all the published experimental reports during this period which provided statistical treatment of the data. Inspection of this table reveals that a majority (61 percent) of the outside replications report significant results ($p < .01$) and that the proportion of significant studies was not significantly greater for the Duke University group ($\chi^2 = 1.70$, 1 *df*).

By 1940, the active methodological controversy was over.

The issues raised were, for the most part, legitimate, and investigators modified their procedures to safeguard their results from methodological criticism. The major issues raised since 1940 center on alleged anomalies in probability theory and the hypothesis of widespread investigator fraud.

Spencer Brown (68, 69) has suggested that statistically significant card-guessing studies provide evidence, not of extrasensory modes of communication, but of fundamental defects in probability theory. He makes three criticisms of random number sequences: (a) published random number

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the paradigm are seldom made; anomalies are ignored. When the anomalies become sufficiently persistent that they can no longer be ignored, they are hotly disputed. Eventually, a new paradigm is tentatively erected which attracts a group of adherents, and a period of crisis ensues which Kuhn calls a paradigm clash.

In this review I have focused at some length on the period of the 1930s, not because it provides the best available evidence for ESP or the best understanding of the processes underlying its operation—it does neither, but rather because it was during this period that the major *substantive methodological issues* were raised and to a large extent consensually resolved. Since 1940, well over 10,000 journal pages devoted to parapsychological research have been published, and at least 250 experimental studies have been reported. The methodological foundations of the research have gradually diversified, enlarging and enriching the scope of inquiry and providing a basis for more sophisticated study. Automated testing equipment has replaced card-guessing in forced-choice ESP tasks, and quantitative methods have been developed for the objective assessment of psi interactions in nonguessing tasks. Psychophysiological techniques, permitting determination of psi-optimal organismic states, have been introduced and utilized in conjunction with experimental methods more closely approximating the conditions under which psi interactions occur *in vivo*. More important, parapsychological investigators have to a large extent shifted their attention away from the "proof-oriented" approach, which can only reaffirm the presence of anomaly, toward systematic attempts to identify the antecedent conditions necessary for the occurrence and detection of psi interactions, the delineation of positive attributes, and the study of individual differences. Only through the pursuit of such "process-oriented" research can we ever hope to achieve the goals of control, assured replicability (or at least predictability), and eventual understanding.

REFERENCES

1. American Psychological Association. "ESP Symposium at the American Psychological Association." *Journal of Parapsychology* 2, 1938, pp. 247–272.
2. Bell, M. "Francis Bacon, Pioneer in Parapsychology." *International Journal of Parapsychology* 6, 1964, pp. 199–208.
3. Bender, H. "The Case of Ilga K.: Report of a Phenomenon of Unusual Perception." *Journal of Parapsychology* 2, 1938, pp. 5–22.
4. Bozarth, J. D., and R. R. Roberts. "Signifying Significant Significance." *American Psychologist* 27, 1972, pp. 774–775.
5. Bridgman, P. W. "Probability, Logic, and ESP." *Science* 123, 1956, pp. 15–17.
6. Camp, B. H. "Statement under 'Notes.'" *Journal of Parapsychology* 1, 1937, p. 305.
7. Carpenter, C. R., and H. R. Phalen. "An Experiment in Card Guessing." *Journal of Parapsychology* 1, 1937, pp. 31–43.
8. Crumbaugh, J. C. An Experimental Study of Extra-sensory Perception. M.A. Thesis, 1938, Southern Methodist University Library.
9. Feller, W. "Statistical Aspects of ESP." *Journal of Parapsychology* 4, 1940, pp. 271–298.
10. Fisher, R. A. Letter to J. B. Rhine, cited in (50), p. 45.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

11. Gibson, E. P. "A Study of Comparative Performance in Several ESP Procedures." *Journal of Parapsychology* 1, 1937, pp. 264–275.
12. Greenwood, J. A. Analysis of a Large Chance Control Series of ESP Data." *Journal of Parapsychology* 2, 1938, pp. 138–146.
13. Gurney, E., F. Myers, and F. Podmore. *Phantasms of the Living* (2 vols.). London: Trubner and Co., 1886.
14. Hansel, C. E. M. "Experiments on Telepathy in Children." *British Journal of Statistical Psychology* 13, 1960. pp. 175–178.
15. Hansel, C. E. M. *ESP—A Scientific Evaluation*. New York: Scribners, 1966.
16. Hansen, F. C. C., and A. Lehmann. "Über Unwillkürliches Flüstern." *Philosophical Studies* 17, 1895, pp. 471–530.
17. Heinlein, C. P., and J. H. Heinlein. "Critique of the Premises and Statistical Methodology of Parapsychology." *Journal of Psychology* 5, 1938, pp. 135–148.
18. Herbert, C. V. C. "Experiment in Extra-sensory Perception." *Journal of the Society for Psychical Research* 30, 1938, pp. 215–218.
19. Humphrey, B. M., and J. A. Clark. "A Comparison of Clairvoyant and Chance Matching." *Journal of Parapsychology* 2, 1938, pp. 31–37.
20. Huntington, E. V. "Is It Chance or ESP?" *American Scholar* 7, 1938, pp. 201–210.
21. Kellogg, C. E. "Dr. J. B. Rhine and Extra-sensory Perception." *Journal of Abnormal and Social Psychology* 31, 1936, pp. 190–193.
22. Kellogg, C. E. "New Evidence (?) for Extra-sensory Perception." *Scientific Monthly*, 1937, pp. 331–341.
23. Kellogg, C. E. Critical comments in (43), pp. 228–238.
24. Kennedy, J. L. "A Methodological Review of Extra-sensory Perception. *Psychological Bulletin* 36, 1939, pp. 59–103.
25. Kennedy, J. L. "The Visual Cues from the Backs of the ESP Cards." *Journal of Psychology* 5, 1938, pp. 149–153.
26. Kennedy, J. L., and H. F. Uphoff. "Experiments on the Nature of Extra-sensory Perception: III. The Recording Error Criticism of Extra-chance Scores." *Journal of Parapsychology* 3, 1939, pp. 226–245.
27. Kennedy, J. L. Critical comments in (43), pp. 218–220.
28. Kuhn, T. S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.
29. Lemmon, V. W. "The Role of Selection in ESP Data." *Journal of Parapsychology* 3, 1939, pp. 104–106.
30. Lemmon, V. W. Critical comments in (43), pp. 222–228.
31. Leuba, C. "An Experiment to Test the Role of Chance in ESP Research." *Journal of Parapsychology* 2, 1938, pp. 217–221.
32. MacFarland, J. D. "Discrimination Shown Between Experimenters by Subjects." *Journal of Parapsychology* 2, 1938, pp. 160–170.
33. MacFarland, J. D., and R. W. George. "Extra-sensory Perception of Normal and Distorted Symbols." *Journal of Parapsychology* 1, 1937, pp. 93–101.
34. Martin, D. R. "Chance and Extra-chance Results in Card-matching." *Journal of Parapsychology* 1, 1937, pp. 185–190.
35. Martin, D. R., and F. P. Stribic. "Studies in Extra-sensory Perception: I. An Analysis of 25,000 Trials; and II. An Analysis of a Second Series of 25,000 Trials." *Journal of Parapsychology* 2, 1938, pp. 23–30; 287–295.
36. Meehl, P. E., and M. Scriven. "Compatibility of Science and ESP." *Science* 123, 1956, pp. 14–15.
37. Murphy, G. "On Limits of Recording Errors." In (1), pp. 262–266.
38. Murphy, G., and E. Taves. "Covariance Methods in the Comparison of Extra-sensory Tasks." *Journal of Parapsychology* 3, 1939, pp. 38–78.
39. Oram, A. T. "An Experiment with Random Numbers." *Journal of the Society for Psychical Research* 37, 1954, pp. 369–377.
40. Oram, A. T. Correspondence. *Journal of the Society for Psychical Research* 38, 1955 pp. 143–144.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

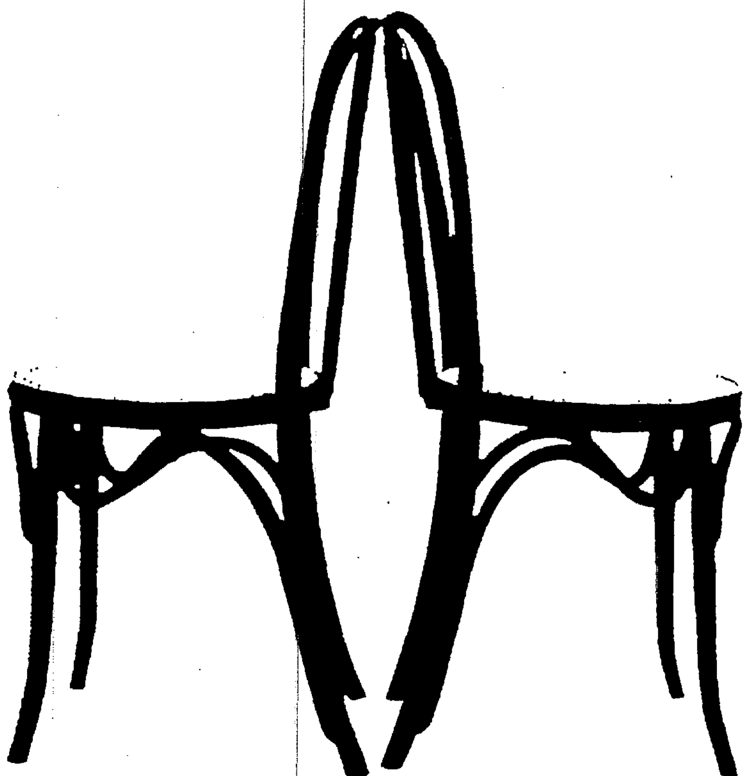
Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

41. Pratt, J. G. "Clairvoyant Blind Matching." *Journal of Parapsychology* 1, 1937, pp. 10–17; and "The Work of Dr. C. Hilton Rice in Extra-sensory Perception, pp. 239–259.
42. Pratt, J. G., and M. Price. "The Experimenter-subject Relationship in Tests for ESP." *Journal of Parapsychology* 2, 1938, pp. 84–94.
43. Pratt, J. G., J. B. Rhine, B. M. Smith, C. E. Stuart, and J. A. Greenwood. *Extrasensory Perception After Sixty Years.* Boston: Bruce Humphries, 1966. (Originally published in 1940 by Henry Holt, New York.)
44. Pratt, J. G., and J. L. Woodruff. "Size of Stimulus Symbols in Extra-sensory Perception." *Journal of Parapsychology* 3, 1939, pp. 121–158.
45. Price, M. "A Comparison of Blind and Seeing Subjects in ESP Tests." *Journal of Parapsychology* 2, 1938, pp. 273–286.
46. Price, M., and M. Pegram. "Extrasensory Perception Among the Blind." *Journal of Parapsychology* 1, 1937, pp. 143–155.
47. Price, G. R. "Science and the Supernatural." *Science* 122, 1955, pp. 359–367.
48. Price, G. R. "Where is the Definitive Experiment?" *Science* 123, 1956, pp. 17–18.
49. Price, G. R. "Apology to Rhine and Soal." *Science* 175, 1972, pp. 359.
50. Rhine, J. B. *Extra-sensory Perception.* Boston: Bruce Humphries, 1964. (Originally published in 1934.)
51. Rhine, J. B. "Telepathy and Clairvoyance in the Normal and Trance States of a Medium." *Character and Personality* 3, 1934, pp. 91–111.
52. Rhine, J. B. "Some Selected Experiments in Extra-sensory Perception." *Journal of Abnormal and Social Psychology* 31, 1936, pp. 216–228.
53. Rhine, J. B. "The Question of Sensory Cues and the Evidence." *Journal of Parapsychology* 1, 1937, pp. 276–291.
54. Rhine, J. B. "ESP Tests with Enclosed Cards." *Journal of Parapsychology* 2, 1938, pp. 199–216.
55. Rhine, J. B. "Comments on 'Science and the Supernatural.'" *Science* 123, 1956, pp. 11–14.
56. Rhine, J. B. "The Experiment Should Fit the Hypothesis." *Science* 123, 1956, p. 19.
57. Rhine, J. B. Comments. *Journal of Parapsychology,* 1974, p. 38.
58. Rhine, J. B., and J. G. Pratt. "A Review of the Pearce-Pratt Distance Series of ESP Tests." *Journal of Parapsychology* 18, 1954, pp. 165–177.
59. Rhine, J. B., B. M. Smith, and J. L. Woodruff. "Experiments Bearing on the Precognition Hypothesis: II. The Role of ESP in the Shuffling of Cards." *Journal of Parapsychology* 2, 1938, pp. 119–131.
60. Riess, B. F. "A Case of High Scores in Card Guessing at a Distance." *Journal of Parapsychology* 1, 1937, pp. 260–263.
61. Shulman, R. "A Study of Card-guessing in Psychotic Subjects." *Journal of Parapsychology* 2, 1938, pp. 96–107.
62. Sinclair, U. *Mental Radio, with a Preface by Albert Einstein.* Springfield, Ill.: Charles C. Thomas, 1962. (Originally published in 1930.)
63. Shapiro, A. "Review of ESP—A Scientific Evaluation by C. E. M. Hansel." *International Journal of Clinical and Experimental Hypnosis* 16, 1968, pp. 133–134.
64. Scott, C. "G. Spencer Brown and Probability: A Critique." *Journal of the Society for Psychical Research* 39, 1958, pp. 217–234.
65. Soal, S. G. "On 'Science and the Supernatural.'" *Science* 123, 1956, pp. 9-11.
66. Soal, S. G., and F. Bateman. *Modern Experiments in Telepathy.* New Haven: Yale University Press, 1954.
67. Soal, S. G., and H. T. Bowden. *The Mind Readers: Recent Experiments in Telepathy.* New Haven, Conn.: Yale University Press, 1954.
68. Spencer Brown, G. "Statistical Significance in Psychical Research." *Nature* 172, 1953, pp. 154–156.
69. Spencer Brown, G. *Probability and Scientific Inference.* New York: Longmans, Green, 1957.
70. Sterling, T. C. "Publication Decisions and Their Possible Effects of Inference Drawn

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

from Tests of Significance—Or Vice Versa." *Journal of the American Statistical sociation* 54, 1959, pp. 30–34.
71. Stuart, C. E. "The Effect of Rate of Movement in Card Matching Tests of Ex sensory Perception." *Journal of Parapsychology* 2, 1938, pp. 171–183.
72. Stuart, C. E. "A Review of Certain Proposed Hypotheses Alternative to Extra-sens Perception." *Journal of Abnormal and Social Psychology* 33, 1938, pp. 57–70.
73. Stuart, C. E., and J. A. Greenwood. "A Review of Criticisms of the Mathemat Evaluation of ESP Data." *Journal of Parapsychology* 1, 1937, pp. 295–304.
74. *Time*, April 29, 1974.
75. Tyrrell, G. N. M. "Some Experiments in Undifferentiated Extra-sensory Percepti *Journal of the Society for Psychical Research* 29, 1935, pp. 52–68.
76. Warner, L. "The Role of Luck in ESP Data." *Journal of Parapsychology* 1, 1937, 84–92.
77. Warner, L. "A Test Case." *Journal of Parapsychology* 1, 1937, pp. 234–238.
78. Willoughby, R. "Critical Comment: The Use of the Probable Error in Evalua Clairvoyance." *Character and Personality* 4, 1935, pp. 79–80.
79. Willoughby, R. "A Critique of Rhine's 'Extra-sensory Perception.'" *Journal of normal and Social Psychology* 30, 1935, pp. 199–207.
80. Willoughby, R. Critical comments in (43), pp. 220–221.
81. Wolins, L. "Needed: Publication of Negative Results." *American Psychologist* 14, 1 p. 598.
82. Wolfle, D. L. "A Review of the Work on Extra-sensory Perception." *American Jou of Psychiatry*, 1938, pp. 943–955.



Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Journal of Parapsychology, Vol. 53, December 1989

"FUTURE TELLING": A META-ANALYSIS OF FORCED-CHOICE PRECOGNITION EXPERIMENTS, 1935–1987

BY CHARLES HONORTON AND DIANE C. FERRARI

ABSTRACT: We report a meta-analysis of forced-choice precognition experiments published in the English-language parapsychological literature between 1935 and 1987. These studies involve attempts by subjects to predict the identity of target stimuli selected randomly over intervals ranging from several hundred milliseconds to one year following the subjects' responses. We retrieved 309 studies reported by 62 investigators. Nearly two million individual trials were contributed by more than 50,000 subjects. Study outcomes are assessed by overall level of statistical significance and effect size. There is a small, but reliable overall effect ($z = 11.41$, $p = 6.3 \times 10^{-25}$). Thirty percent of the studies (by 40 investigators) are significant at the 5% significance level. Assessment of vulnerability to selective reporting indicates that a ratio of 46 unreported studies averaging null results would be required for each reported study in order to reduce the overall result to nonsignificance. No systematic relationship was found between study outcomes and eight indices of research quality. Effect size has remained essentially constant over the survey period, whereas research quality has improved substantially. Four moderating variables appear to covary significantly with study outcome: Studies using subjects selected on the basis of prior testing performance show significantly larger effects than studies using unselected subjects. Subjects tested individually by an experimenter show significantly larger effects than those tested in groups. Studies in which subjects are given trial-by-trial or run-score feedback have significantly larger effects than those with delayed or no subject feedback. Studies with brief intervals between subjects' responses and target generation show significantly stronger effects than studies involving longer intervals. The combined impact of these moderating variables appears to be very strong. Independently significant outcomes are observed in seven of the eight studies using selected subjects, who were tested individually and received trial-by-trial feedback.

Precognition refers to the noninferential prediction of future events. Anecdotal claims of "future telling" have occurred throughout human history in virtually every culture and period. Today such

This work was funded by SRI International and the John E. Fetzer Foundation. We wish to thank our PRL colleague George P. Hansen, who is primarily responsible for retrieving the studies used in the meta-analysis. We are grateful to Edwin C. May, Jessica Utts, and to five anonymous reviewers at SRI for valuable comments on an earlier draft of this report. Valuable comments were also made by Ephraim Schechter and by three anonymous referees. The division of authorship responsibility is as follows: Honorton is responsible for the design of the meta-analysis, definition of study coding criteria, the actual analyses, and the report itself. Ferrari coded the individual research reports in consultation with Honorton and/or Hansen.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

claims are generally believed to be based on factors such as delusion, irrationality, and superstitious thinking. The concept of precognition runs counter to accepted notions of causality and appears to conflict with current scientific theory. Nevertheless, over the past half-century a substantial number of experiments have been reported claiming empirical support for the hypothesis of precognition. Subjects in forced-choice experiments, according to many reports, have correctly predicted to a statistically significant degree the identity (or order) of target stimuli randomly selected at a later time.

We performed a meta-analysis of forced-choice precognition experiments published in the English-language research literature between 1935 and 1987. Four major questions were addressed through this meta-analysis: (1) Is there overall evidence for accurate target identification (above-chance hitting) in experimental precognition studies? (2) What is the magnitude of the overall precognition effect? (3) Is the observed effect related to variations in methodological quality that could allow a more conventional explanation? (4) Does precognition performance vary systematically with potential moderating variables, such as differences in subject populations, stimulus conditions, experimental setting, knowledge of results, and time interval between subject response and target generation?

Delineating the Domain

Retrieval of Studies

Parapsychological research is still academically taboo, and it is unlikely that there have been many dissertations and theses in this area that have escaped publication. Our retrieval of studies for this meta-analysis is therefore based on the published literature. The studies include all forced-choice precognition experiments appearing in the peer-reviewed English-language parapsychology journals: *Journal of Parapsychology, Journal* (and *Proceedings) of the Society for Psychical Research, Journal of the American Society for Psychical Research, European Journal of Parapsychology* (including the *Research Letter* of the Utrecht University Parapsychology Laboratory), and abstracts of peer-reviewed papers presented at Parapsychological Association meetings published in *Research in Parapsychology.*

Criteria for Inclusion

Our review is restricted to fixed-length studies in which significance levels and effect sizes based on direct hitting can be calculated. Studies using outcome variables other than direct hitting, such as run-score variance and displacement effects, are included only if the report provides relevant information on direct hits (i.e., number of trials, hits, and probability of a hit). Finally, we exclude studies conducted by two investigators, S. G. Soal and Walter J. Levy, whose work has been unreliable.

Many published reports contain more than one experiment or experimental unit. In experiments involving multiple conditions, significance levels and effect sizes are calculated for each condition.

Outcome Measures

Significance level. Significance levels (z scores) were calculated for each study from the reported number of trials, hits, and probability of success using the normal approximation to the binomial distribution with continuity correction. Positive *z* scores indicate above chance scoring, and negative *z* scores reflect below-chance scoring.

Effect size. Because most parapsychological experiments, particularly those in the older literature, have used the trial rather than the subject as the sampling unit, we use a trial-based estimator of effect size. The effect size (*ES*) for each study is the *z* score divided by the square root of the number of trials in the study.[1]

General Characteristics of the Domain

We located 309 studies in 113 separate publications. These studies were contributed by 62 different senior authors and were published over a 53-year period, between 1935 and 1987. Considering the half-century time-span over which the precognition experiments were conducted, it is not surprising that the studies are very diverse.

The database comprises nearly two million individual trials and more than 50,000 subjects. Study sample sizes range from 25 to 297,060 trials (median = 1,194). The number of subjects ranges from 1 to 29,706 (median = 16). The studies use a variety of methodologies, ranging from guessing ESP cards and other card symbols to automated random number generator experiments. The domain encompasses diverse subject populations: the most frequently used

[1] Elsewhere (Honorton, 1985), we have used the effect size index Cohen's *h* (Cohen, 1977), and one referee has asked that we explain why we are now using $z/N^{1/2}$. The answer is that *h* and $z/N^{1/2}$ yield virtually identical results, and $z/N^{1/2}$ is computationally simpler. For the present sample of 309 precognition studies, the mean difference between the two indices is .00047, and the standard deviation of the difference is .026: $t(308) = 0.312$, $p = .756$, two-tailed. The correlation between the two indices is .97.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4
Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 1
OVERALL SIGNIFICANCE LEVEL AND EFFECT SIZE

	z	ES
Mean	0.65	0.020
SD	2.68	0.100
Lower 95% confidence estimate	0.40	0.011

Combined $z = 11.41$, $p = 6.3 \times 10^{-25}$

"Fail-safe N" = 14,268

$t(ES) = 3.51$, 308 df, $p = .00025$

population is students (in approximately 40% of the studies); the least frequently used populations are the experimenters themselves and animals (each used in about 5% of the studies).

Though a few studies tested subjects through the mail, more typically subjects were tested in person, either individually or in groups. Target selection methods included no randomization at all (studies using "quasi-random" naturalistic events), informal methods including manual card-shuffling or dice-throwing, and formal methods, primarily random number tables or random number generators. The time interval between the subjects' responses and target generation varied from less than one second to one year.

OVERALL CUMULATION

Evidence for an overall effect is strong. As shown in the top part of Table 1, the overall results are highly significant.[2] Lower bound (one-tailed) 95% confidence estimates of the mean z score and ES are displayed in the bottom portion of Table 1.

Ninety-two studies (30%) show significant hitting at the 5% level, and significant outcomes are contributed by 40 different investigators. The z scores correlate significantly with sample size: $r(307) = .156$, $p = .003$. The mean number of trials for significant studies is 34% larger than the mean number of trials for nonsignificant studies.

[2] The statistical analyses presented here were performed using SYSTAT (Wilkinson, 1988). When t tests are reported on samples with unequal variances, they are calculated using the separate variances within groups for the error and degrees of freedom following Brownlee (1965). Unless otherwise specified, p levels are one-tailed. Combined z's are based on Stouffer's method (Rosenthal, 1984).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

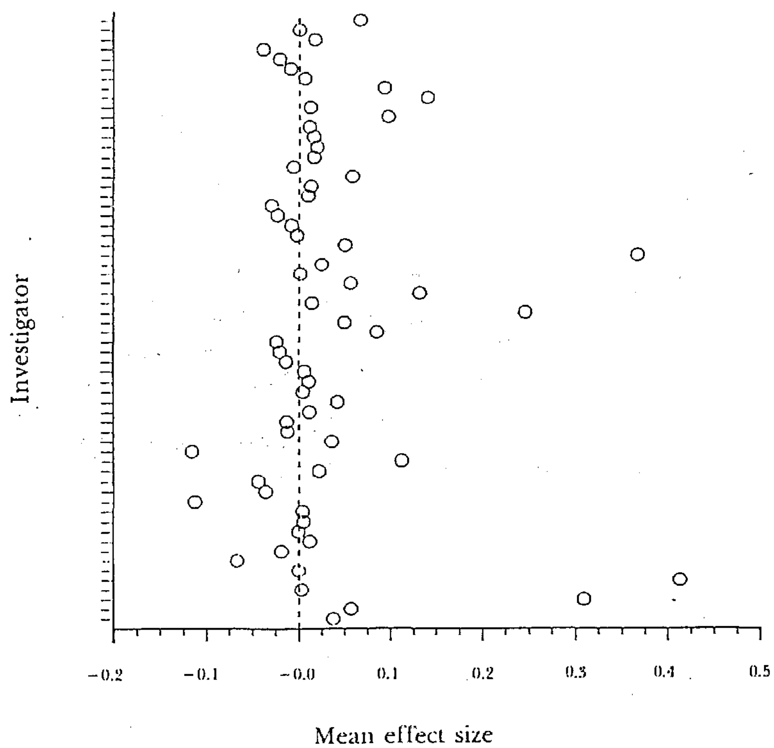


Figure 1. Mean effect size by investigator. $N = 62$ investigators.

Replication Across Investigators

Virtually the same picture emerges when the cumulation is by investigator rather than study as the unit of analysis; the combined z is 12.13, and 23 of the 62 investigators (37%) have overall outcomes significant at the 5% level. The mean (investigator) effect size is 0.033 ($SD = .093$).

There is a significant difference in the mean ES across investigators, but it is surprisingly small: Kruskal-Wallis one-way ANOVA by ranks, $\chi^2(61) = 82.71$, $p = .034$. The effect is clearly not due to a few major contributors. If investigators contributing more than three studies are eliminated, leaving 33 investigators, the combined z is still 6.00 ($p = 1.25 \times 10^{-9}$) and the mean ES is .028 ($SD = .091$). Figure 1 shows the mean effect sizes by investigator.

These results indicate substantial cross-investigator replicability and directly contradict the claim of critics such as Akers (1987) that successful parapsychological outcomes are achieved by only a few investigators.

The Filedrawer Problem

A well-known reporting bias exists throughout the behavioral sciences favoring publication of "significant" studies (e.g., Sterling, 1959). The extreme view of this "filedrawer problem" is that "the journals are filled with the 5% of the studies that show Type I errors, while the filedrawers back at the lab are filled with the 95% of the studies that show nonsignificance..." (Rosenthal, 1984, p. 108). Recognizing the importance of this problem, the Parapsychological Association in 1975 adopted an official policy against selective reporting of positive results.[3] Examination of the parapsychological literature shows that nonsignificant results are frequently published, and, in the precognition database, 70% of the studies have reported nonsignificant results. Nevertheless, 75% of the precognition studies were published before 1975, and we must ask to what extent selective publication bias could account for the cumulative effects we observe.

The central section of Table 1 uses Rosenthal's (1984) "fail-safe N" statistic to estimate the number of unreported studies with z scores averaging zero that would be necessary to reduce the known database to nonsignificance. The filedrawer estimate indicates that over 46 unreported studies must exist for each reported study to reduce the cumulative outcome to a nonsignificant level.

A different approach to the filedrawer problem is described by Dawes, Landman, and Williams (1984; personal communication from Dawes to Honorton, July 14, 1988). Their truncated normal curve analysis, like Rosenthal's "fail-safe N," is based on normal curve assumptions. Their null hypothesis is that z scores above some critical level (e.g., $z = 1.65$, 1.96, etc.) are randomly sampled from $N(0,1)$ above that critical level. The alternative to the null hypothesis is that, because there is some real effect, the distribution of z's is shifted to the right of 0 and the z's will be larger than predicted by the null. For a critical level of $z = 1.65$, the expected mean z is 2.06 and the variance is .14. In the precognition database, there are 92 studies with z's > 1.65. Their average is 3.61, not 2.06 as predicted by the null hypothesis. Since the variance of the normal truncated above 1.65 is .14, the *test* z (using the Central Limit Theorem) comparing 3.61 to 2.06 is 39.84 [1.55 divided by $(.14/92)^{1/2}$]. Here, p is virtually zero. Similar results are found with cut points of 1.96, 2.33, and 2.58.

On the basis of these analyses, we conclude that the cumulative significance of the precognition studies cannot satisfactorily be explained by selective reporting.

[3] Analyses indicate no significant difference in the magnitude of reported study outcomes before and after 1975. The mean *ES* for studies prior to 1975 is 0.021 (SD = .099), and for studies reported thereafter the mean is 0.017 (SD = .106): $t(307) = 0.28$, $p = .782$, two-tailed.

Outlier Reduction

Although the overall z scores and effect sizes cannot reasonably be attributed to chance, inspection of the standard deviations in Table 1 indicates that the study outcomes are extremely heterogeneous. Given the diversity of methods, subject populations, and other study features that characterize this research domain, this is not surprising.

The study outcomes are in fact extremely heterogeneous. Although a major objective of this meta-analysis is to account for the variability across studies by blocking on differences in study quality, procedural features, and sampling characteristics, the database clearly contains extreme outliers. The z scores range from -5.1 to 19.6, a 25-sigma spread! The standardized index of kurtosis (g_2) is 9.47, suggesting that the tails of the distribution are much too long for a normal distribution.

We eliminated the extreme outliers by performing a "10 percent trim" on the study z scores (Barnett & Lewis, 1978). This involves eliminating studies with z scores in the upper and lower 10% of the distribution, and results in an adjusted sample of 248 studies. The trimmed z scores range from -2.24 to 3.21 ($g_2 = -1.1$). The revised z scores and effect sizes are presented in Table 2.

Elimination of extreme outliers reduces the combined z scores by approximately one half, but the outcomes remain highly significant. Twenty-five percent of the studies (62/248) show overall significant hitting at the 5% level. Lower bound confidence estimates show that the mean z's and effect sizes are above 0 at the 95% confidence level.

Elimination of outliers reduces the total number of investigators from 62 to 57, but the results remain basically the same when the analyses are based on investigators rather than studies. The combined z is 6.84; 18 of the 57 investigators (31.6%) have overall sig-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4
Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 2
SIGNIFICANCE LEVEL AND EFFECT SIZE FOR TRIMMED SAMPLE

	z	*ES*
Mean	0.38	0.012
SD	1.45	0.065
Lower 95% confidence estimate	0.23	0.005

Combined $z = 6.02$, $p = 1.1 \times 10^{-9}$
$t(ES) = 2.90$, 247 *df*, $p = .002$

nificant outcomes at the 5% level. The mean (investigator) *ES* is 0.0[illegible]0 (*SD* = .05).

For the trimmed sample, the difference in *ES* across investigators is *not* significant: Kruskal-Wallis one-way ANOVA by ranks, $\chi^2(56) = 59.34$, $p = .355$. If investigators contributing more than three studies are eliminated, leaving 37 investigators, the combined z is still 5.00 ($p = 3.0 \times 10^{-7}$) and the mean *ES* is 0.022 (*SD* = .05[illegible]). Figure 2 shows the mean effect size by investigator.

Thus, elimination of the outliers does not substantially affect the conclusions drawn from our analysis of the database as a whole. There clearly is a nonchance effect. In the remainder of this report, we use the trimmed sample to examine covariations in effect size and a variety of methodological and other study features.

STUDY QUALITY

Because target stimuli in precognition experiments are selected only after the subjects' responses have been registered, precognition studies are usually not vulnerable to sensory leakage problems. Other potential threats to validity must, however, be considered. The problem of variations in research quality remains a source of controversy in meta-analysis. Some meta-analysts advocate eliminating low quality studies whereas others recommend empirically accessing the impact of variations in quality on study outcome. Rosenthal (1984) points out that the practice of discarding studies is equivalent to assigning them weights of zero, and he recommends weighting study z scores in relation to ratings of research quality.

Study Quality Criteria

Ideally, the assessment of study quality should be performed by knowledgeable specialists who are blind to the study outcomes. In

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

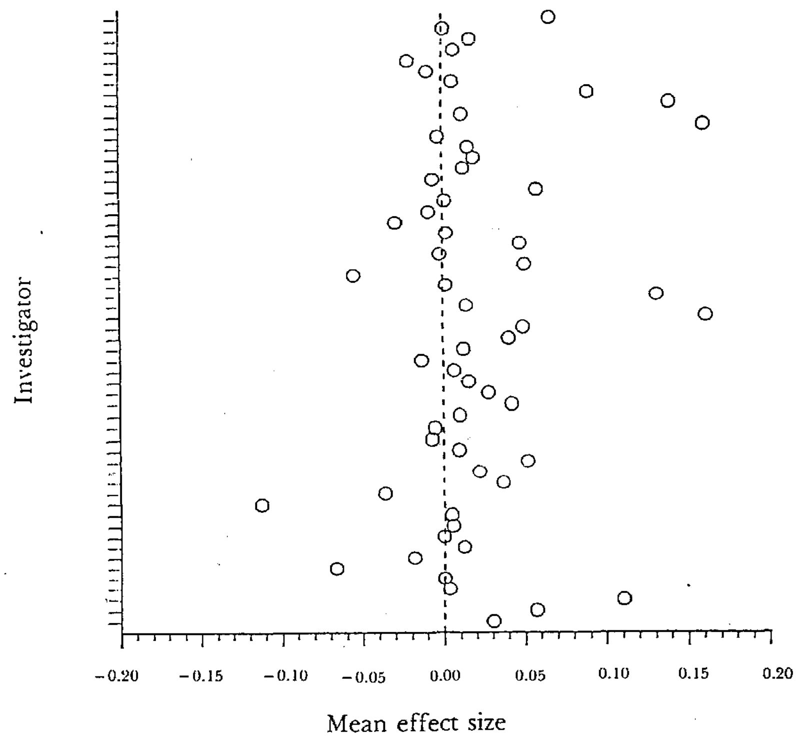


Figure 2. Mean effect size by investigator for trimmed sample. N = 57 investigators.

practice, this is usually not feasible, particularly when, as in the present case, large numbers of studies are involved. For our analysis of study quality, statistical and methodological variables are defined and coded in terms of procedural descriptions (or their absence) in the research reports. This approach was used in an earlier meta-analysis of psi ganzfeld research (Honorton, 1985), and it led to study quality ratings that were generally in agreement, $r(26) = .766$, $p = 10^{-6}$, with independent "flaw" ratings by an outside critic (Hyman, 1985).

One point is given (or withheld) for each of the following eight criteria:

Specification of sample size. Does the investigator preplan the number of trials to be included in the study or is the study vulnerable to the possibility of optional stopping? Credit is given to reports that explicitly specify the sample size. Studies involving group testing, in which it is not feasible to specify the sample size precisely, are also

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

given credit. No credit is given to studies in which the sample size is either not preplanned or not addressed in the experimental report.

Preplanned analysis. Is the method of statistical analysis, including the outcome (dependent variable) measure, preplanned? Credit is given to studies explicitly specifying the form of analysis and the outcome measure. No credit is given to those not explicitly stating the form of the analysis or those in which the analysis is clearly post hoc.

Randomization method. Credit is given for use of random number tables, random number generators, and mechanical shufflers. No credit is given for failure to randomize (i.e., use of "quasi-random naturalistic events") or for informal methods such as hand-shuffling, die-casting, and drawing lots.

Controls. Credit is given to studies reporting randomness control checks, such as random number generator (RNG) control series and empirical cross-check controls.

Recording. One point is allotted for automated recording of targets and responses, and another for duplicate recording.

Checking. One point is allotted for automated checking of matches between target and response, and another for duplicate checking of hits.

Study Quality Analysis

Each study received a quality weight between 0 and 8 (mean = 3.3, SD = 1.8). We find no significant relationship between study quality and ES: $r(246)$ = .081, p = .202, two-tailed. This tendency for study outcomes to correlate *positively* with study quality has the consequence that the quality-weighted z score of 6.26 is slightly *larger* than the unweighted z of 6.02. Table 3 shows the correlations between effect size and each of the eight individual quality measures.[4] The mean effect sizes by quality level are displayed graphically in Figure 3.

[4] The correlation between ES and study quality is also nonsignificant for the untrimmed sample of 309 studies: $r(307) = -.060$, p = .289. The quality-weighted z score is 7.38: $p = 2.32 \times 10^{-13}$. However, three of the individual quality measures are significantly related to performance. Controls and duplicate checking correlate significantly positively with ES, and randomization correlates significantly negatively with ES. These correlations appear to be due to a few studies with z scores that are extreme outliers ($z > 7$). When the 10 studies with $z > 7$ are eliminated, the significant correlations between quality and ES disappear.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 3
CORRELATIONS BETWEEN EFFECT SIZE AND QUALITY MEASURES

Quality measure	$r(246)$
Sample size specified in advance	− .100
Preplanned analysis	− .001
Randomization	− .011
Controls	.058
Automated recording	.169
Duplicate recording	.047
Automated checking	.136
Duplicate checking	.078

Quality Extremes

Is there a tendency for extremely weak studies to show larger effects than exceptionally "good" studies? Analysis on the extremes of the quality ratings indicates that this is not the case.

This analysis, based on the untrimmed sample of 309 studies, uses studies with quality ratings outside the interquartile range of the rating distribution (median = 4, Q_1 = 2, Q_3 = 5). There are 56 "low-quality" studies (ratings of 0–1) and 35 "high-quality" studies (ratings of 6–8). The high-quality studies have effect sizes that are not significantly lower than the low-quality studies; the ES means are 0.017 (SD = 0.063) and 0.037 (SD = 0.137), for the low- and high-quality studies, respectively: $t(82) = -.92$, p = .358, two-tailed.

Quality Variation in Publication Sources

Precognition ES is not significantly related to source of publication: Kruskal-Wallis one-way ANOVA, $\chi^2(4)$ = 0.78, p = .942. However, the sources of publication differ significantly in study quality: Kruskal-Wallis one-way ANOVA, $\chi^2(4)$ = 17.19, p = .002. This is due largely to the lower quality of studies published in the *Journal of the Society for Psychical Research* and in *Research in Parapsychology*.

Study Quality in Relation to Year of Publication

Precognition effect size has remained constant over a half-century of research, even though the methodological quality of the re-

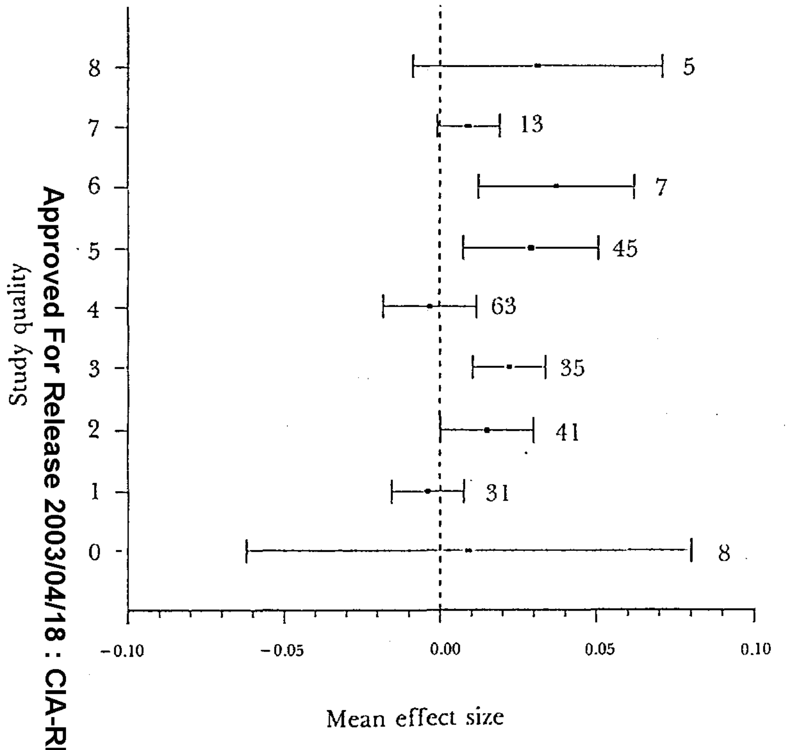


Figure [illegible]. Precognition effect size in relation to study quality, with 95% confidence limits. N = 248 studies.

search has improved significantly during this period. The correlation between *ES* and year of publication is −.071: $t(307) = -1.25$, p = .2[illegible]3, two-tailed. Study quality and year of publication are, however, positively and significantly correlated: $r(246) = .282$, $p = 2 \times 10^{-7}$, two-tailed.

Critics of parapsychology have long believed that evidence for parapsychological effects disappears as the methodological rigor increases. The precognition database does not support this belief.

"Real-Time" Alternatives to Precognition

Investigators have long been aware of the possibility that precognition effects could be modeled without assuming either time reversal or backward causality. For example, outcomes from studies with

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

targets based on indeterminate random number generators (RNGs) could be due to a causal influence on the RNG—a psychokinetic (PK) effect—rather than information acquisition concerning its future state. In experiments with targets based on prepared tables of random numbers, the possibility exists that the experimenter or other randomizer may be the actual psi source, unconsciously using "real-time" ESP combined with PK to choose an entry point in the random number sequence that will significantly match the "subject's" responses. While the latter possibility may seem far-fetched, it cannot be logically eliminated if one accepts the existing evidence for contemporaneous ESP and PK, and it has been argued that it is less far-fetched than the alternative of "true" precognition.

Morris (1982) discusses models of experimental precognition based on "real-time" psi alternatives and methods for testing "true" precognition. In general terms, these methods constrain the selection of the target sequence so as to eliminate nonprecognitive psi intervention. In the most common procedure, attributed to Mangan (1955), dice are thrown to generate a set of numbers that are mathematically manipulated to obtain an entry point in the random number table. This procedure is sufficiently complex "as to be apparently beyond the capacities of the human brain, thus ruling out PK because the 'PKer' would not know what to do even via ESP" (Morris, 1982, p. 329).

Two features of precognition study target determination procedures were coded to assess "real-time" psi alternatives to precognition: method of determining random number table entry point and use of Mangan's method.

Methods of eliminating "real-time" psi alternatives have not been used in studies with random number generators and have only been used in a small number of studies involving randomization by hand-shuffling. These analyses are therefore restricted to studies using random number tables (N = 138).

Method of Determining RNT Entry Point

The reports describe six different methods of obtaining entry points in random number tables. If the study outcomes were due to subjects' precognitive functioning rather than to alternative psi modes on the part of the experimenter or the experimenter's assistants, there should be no difference in mean effect size across the various methods used to determine the entry point. Indeed, our analysis indicates that the study effect sizes do not vary systemati-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

cally as a function of method of determining the entry point: Kruskal-Wallis one-way ANOVA by ranks: $\chi^2(5) = 7.32$, $p = .198$.

Use of Mangan's Method

We find no significant difference in *ES* between studies using complex calculations of the type introduced by Mangan to fix the random number table entry point and those that do not use such calculations: $t(45) = 0.38$, $p = .370$, two-tailed.

Moderating Variables

The stability of precognition study outcomes over a 50-year period, which we described earlier, is also bad news. It shows that investigators in this area have yet to develop sufficient understanding of the conditions underlying the occurrence (or detection) of these effects to reliably increase their magnitude. We have identified four variables that appear to covary systematically with precognition *ES*: (1) selected versus unselected subjects, (2) individual versus group testing, (3) feedback level, and (4) time interval between subject response and target generation.

The analyses use the raw study *z* scores and effect sizes; we found that this results in uniformly more conservative estimates of relationships with moderating variables than when the analyses are based on quality-weighted *z* scores and effect sizes.

Selected Versus Unselected Subjects

Our meta-analysis identifies eight subject populations: unspecified subject populations, mixtures of several different populations, animals, students, children, "volunteers," experimenter(s), and selected subjects.

Effect size magnitude does not vary significantly across these eight subject populations: Kruskal-Wallis one-way ANOVA, $\chi^2(7) = 10.90$, $p = .143$. Effect sizes by subject population are displayed in Figure 4.

However, studies using subjects selected on the basis of prior performance in experiments or pilot tests show significantly larger effects than studies using unselected subjects. As shown in Table 4, 60% of the studies with selected subjects are significant at the 5% level. The mean *z* score for these studies is 1.39 ($SD = 1.40$). The *ES* is significantly higher for selected-subjects studies than for studies with unselected subjects. The *t* test of the difference in mean *ES* is equivalent to a point-biserial correlation of .198.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

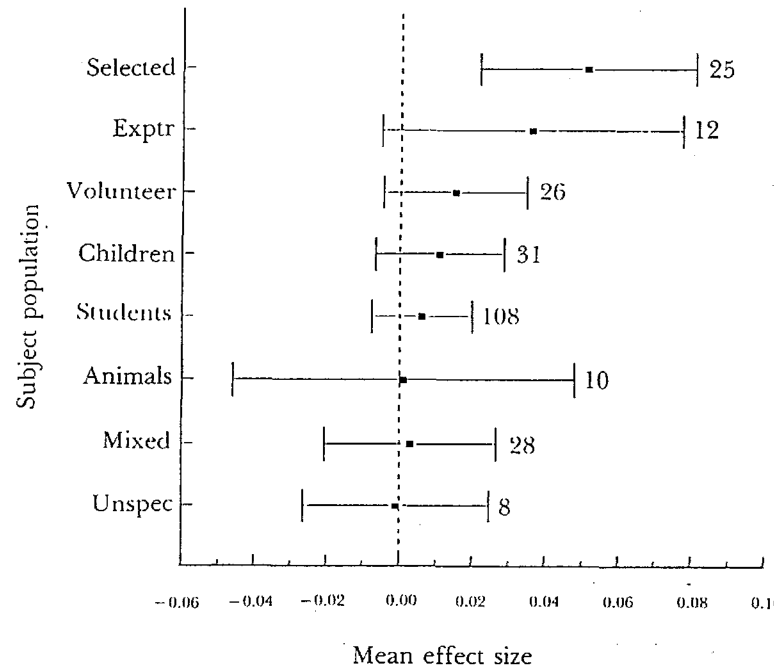


Figure 4. Precognition effect size by subject population, with 95% confidence limits. $N = 248$ studies.

Does this difference result from less stringent controls in studies with selected subjects? The answer appears to be "No." The average quality of studies with selected subjects is *higher* than studies using

Table 4
Selected Versus Unselected Subjects

	Selected	Unselected
N studies	25	223
Combined *z*	6.89	4.04
Studies with $p < .05$	60%	21%
Mean *ES*	.051	.008
SD_{ES}	.075	.063

$t(246) = 3.16$, $p = .001$

TABLE 5
INDIVIDUAL VERSUS GROUP TESTING

	Individual	Group
N studies	97	105
Combined *z*	6.64	1.29
Studies with $p < .05$	30%	19%
Mean *ES*	.021	.004
SD_{ES}	.060	.066

$t\ (200) = 1.89,\ p = .03$

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

unselected subjects: $t(27) = 1.51$, $p = .142$, two-tailed. This result appears to reflect a general tendency toward increased rigor and more detailed reporting in studies with selected subjects.

Individual Versus Group Testing

Subjects were tested in groups, individually, or through the mail. Studies in which subjects were tested individually by an experimenter have a significantly larger mean *ES* than studies involving group testing (Table 5).

The *t* test of the difference is equivalent to a point-biserial correlation of .132, favoring individual testing. Of the studies with subjects tested individually, 30% are significant at the 5% level.

The methodological quality of studies with subjects tested individually is significantly higher than that of studies involving group testing: $t(137) = 3.08$, $p = .003$, two-tailed. This result is consistent with the conjecture that group experiments are frequently conducted as "targets of opportunity" and may often be carried out hastily in an afternoon without the preparation and planning that go into a study with individual subjects that may be conducted over a period of weeks or months.

Thirty-five studies were conducted through the mail. In these studies, subjects completed the task at their leisure and mailed their responses to the investigator. These correspondence studies yield outcomes similar to those involving individual testing. The combined *z* score is 2.66, with a mean *ES* of 0.018 ($SD = .082$). Ten correspondence studies (25.7%) are significant at the 5% level.

Eleven studies are unclassifiable with regard to experimental setting.

TABLE 6
FEEDBACK RECEIVED BY SUBJECTS

	Feedback of Results			
	None	Delayed	Run score	Trial-by-trial
N studies	15	21	21	47
Combined *z*	−1.30	2.11	4.74	6.98
Studies with $p < .05$	0.0%	19.0%	33.3%	42.6%
Mean *ES*	−.001	.009	.023	.035
SD_{ES}	.028	.036	.048	.072

Feedback

A significant positive relationship exists between the degree of feedback subjects receive about their performance and precognitive effect size (Table 6).

Subject feedback information is available for 104 studies. These studies fall into four feedback categories: no feedback, delayed feedback (usually notification by mail), run-score feedback, and trial-by-trial feedback. We gave these categories numerical values between 0 and 3. Precognition effect size correlates .231 with feedback level (102 *df*, $p = .009$). Of the 47 studies involving trial-by-trial feedback, 20 (42.6%) are significant at the 5% level. None of the studies without subject feedback are significant.

Feedback level correlates positively though not significantly with research quality: $r(102) = .173$, $p = .082$, two-tailed. Inadequate randomization is the most plausible source of potential artifacts in studies with trial-by-trial feedback. We performed a separate analysis on the 47 studies in this group. Studies using formal methods of randomization do not differ significantly in mean *ES* from those with informal randomization: $t(15) = 0.67$, $p = .590$, two-tailed. Similarly, studies reporting randomness control data do not differ significantly in *ES* from those not including randomness controls: $t(42) = 0.79$, $p = .436$, two-tailed.

Time Interval

The interval between the subject's response and target selection ranges from less than one second to one year. Information about the time interval is available for 144 studies. This information, how-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

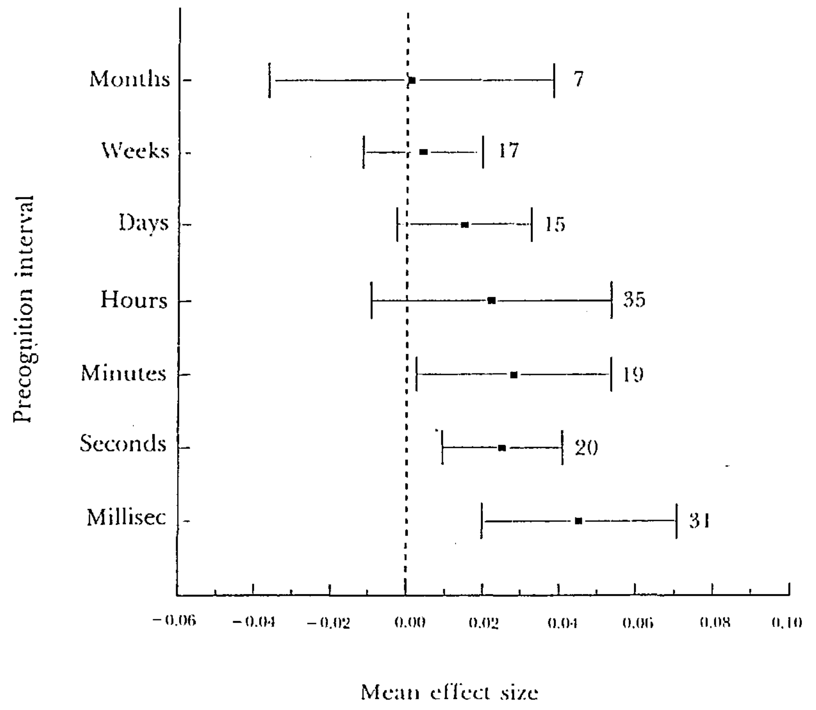


Figure 5. Effect size by precognition interval, with 95% confidence limits. N = 144 studies.

ever, is often imprecise. Our analysis of the relationship between precognitive ES and time interval is therefore limited to seven broad interval categories: milliseconds, seconds, minutes, hours, days, weeks, and months. (Effect sizes by precognition interval are displayed in Figure 5.)

Although it is confounded with degree of feedback, there is a significant decline in precognition ES over increasing temporal distance: $r(142) = -.199$, $p = .017$, two-tailed. The largest effects occur over the millisecond interval: $N = 31$ studies, combined $z = 6.03$, mean $ES = 0.045$, $SD = .073$. The smallest effects occur over periods ranging from a month to a year: $N = 7$, combined $z = 0.53$, mean $ES = 0.001$, $SD = .049$.

Interestingly, the decline of precognition performance over increasing temporal distances results entirely from studies using *unselected* subjects: $r(122) = -.235$, $p = .009$, two-tailed. Studies with selected subjects show a nonsignificant positive relationship between ES and time interval: $r(18) = .077$, $p = .745$, two-tailed. Although the difference between these two correlations is not significant ($z = 1.24$), this suggests that the origin of the decline over time may be motivational rather than the result of some intrinsic physical boundary condition. The relationship between precognition ES and feedback also supports this conjecture. Nevertheless, any finding suggesting potential boundary conditions on the phenomenon should be vigorously pursued.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Influence of Moderating Variables in Combination

The above analyses examine the impact of each moderating variable in isolation. In this final set of analyses, we explore their joint influence on precognition performance. For this purpose, we identify two subgroups of studies. One subgroup is characterized by the use of selected subjects tested individually with trial-by-trial feedback. We refer to this as the *Optimal* group (N = 8 studies). The second group is characterized by the use of unselected subjects tested in groups with no feedback. We refer to this as the *Suboptimal* group (N = 9 studies).

The Optimal studies are contributed by four independent investigators and the Suboptimal studies are contributed by two of the same four investigators. All of the Optimal studies involve short precognition time intervals (millisecond interval); the Suboptimal studies involve longer intervals (intervals of weeks or months). All of the Optimal studies and 5 of the 9 Suboptimal studies use RNG methodology. The two groups do not differ significantly in average sample size. The mean study quality for the Optimal group is significantly higher than that of the Suboptimal studies: Optimal mean = 6.63, $SD = 0.92$; Suboptimal mean = 3.44, $SD = 0.53$; $t(10) = 8.63$, $p = 3.3 \times 10^{-6}$, two-tailed.

The combined impact of the moderating variables appears to be quite strong (Table 7). Seven of the 8 Optimal studies (87.5%) are independently significant at the 5% level, whereas none of the Suboptimal studies are statistically significant. All four investigators contributing studies to the Optimal group have significant outcomes.[5]

---

[5] In the untrimmed sample of 309 studies, there are a total of 17 Optimal studies. The mean ES is 0.117 ($SD = .154$), and the combined z is 15.84. The percentage of independently significant studies is virtually the same as it is in the trimmed sample: 15 of the 17 studies (88.2%) are significant.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 7
IMPACT OF MODERATORS IN COMBINATION

	"Optimal" studies	"Suboptimal" studies
N studies	8	9
Combined *z*	6.14	− 1.29
Studies with $p < .05$	87.5%	0.0%
Mean *ES*	.055	.005
SD_{ES}	.045	.035

$t(15) = 2.61, p = .01$
$r = .559$

These results are quite striking and suggest that future studies combining these moderators should yield especially reliable effects.

SUMMARY AND CONCLUSIONS

Our meta-analysis of forced-choice precognition experiments confirms the existence of a small but highly significant precognition effect. The effect appears to be replicable; significant outcomes are reported by 40 investigators using a variety of methodological paradigms and subject populations.

The precognition effect is statistically very robust: it remains highly significant despite elimination of studies with z scores in the upper and lower 10% of the z-score distribution and when a third of the remaining investigators—the major contributors of precognition studies—are eliminated.

Estimates of the "filedrawer" problem and consideration of parapsychological publication practices indicate that the precognition effect cannot plausibly be explained on the basis of selective publication bias. Analyses of precognition effect sizes in relation to eight measures of research quality fail to support the hypothesis that the observed effect is driven to any appreciable extent by methodological flaws; indeed, several analyses indicate that methodologically superior studies yield stronger effects than methodologically weaker studies.

Analyses of parapsychological alternatives to precognition, although limited to the subset of studies using random number tables, provide no support for the hypothesis that the effect results from

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the operation of contemporaneous ESP and PK at the time of randomization.

Although the overall precognition effect size is small, this does not imply that it has no practical consequences. It is, for example, of the same order of magnitude as effect sizes leading to the early termination of several major medical research studies. In 1981, the National Heart, Lung, and Blood Institute discontinued its study of propranolol because the results were so favorable to the propranolol treatment that it would be unethical to continue placebo treatment (Kolata, 1981); the effect size was 0.04. More recently, The Steering Committee of the Physicians' Health Study Research Group (1988), in a widely publicized report, terminated its study of the effects of aspirin in the prevention of heart attacks for the same reason. The aspirin group suffered significantly fewer heart attacks than a placebo control group; the associated effect size was 0.03.

The most important outcome of the meta-analysis is the identification of several moderating variables that appear to covary systematically with precognition performance. The largest effects are observed in studies using subjects selected on the basis of prior test performance, who are tested individually, and who receive trial-by-trial feedback. The outcomes of studies combining these factors contrast sharply with the null outcomes associated with the combination of group testing, unselected subjects, and no feedback of results. Because the two groups of studies were conducted by a subset of the same investigators, it is unlikely that the observed difference in performance is due to experimenter effects. Indeed, these outcomes underscore the importance of carefully examining differences in subject populations, test setting, and so forth, before resorting to facile "explanations" based on psi-mediated experimenter effects or the "elusiveness of psi."

The identification of these moderating variables has important implications for our understanding of the phenomena and provides clear direction for future research. The existence of moderating variables indicates that the precognition effect is not merely an unexplained departure from a theoretical chance baseline, but rather is an effect that covaries with factors known to influence more familiar aspects of human performance. It should now be possible to exploit these moderating factors to increase the magnitude and reliability of precognition effects in new studies.

REFERENCES

AKERS, C. (1987). Parapsychology is science, but its findings are inconclusive. *Behavioral and Brain Sciences*, **10**, 566–568.

Barnett, V., & Lewis, T. (1978). *Outliers in statistical data.* New York: John Wiley & Sons.
Brownlee, K. A. (1965). *Statistical theory and methodology in science and engineering.* New York: John Wiley & Sons.
Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.
Dawes, R. M., Landman, J., & Williams, J. (1984). Reply to Kurosawa. *American Psychologist,* **39**, 74–75.
Honorton, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. *Journal of Parapsychology,* **49**, 51–92.
Hyman, R. (1985). The ganzfeld psi experiment: A critical appraisal. *Journal of Parapsychology,* **49**, 3–50.
Kolata, G. B. (1981). Drug found to help heart attack survivors. *Science,* **214**, 774–775.
Mangan, G. L. (1955). Evidence of displacement in a precognition test. *Journal of Parapsychology,* **19**, 35–44.
Morris, R. L. (1982). Assessing experimental support for true precognition. *Journal of Parapsychology,* **46**, 321–336.
Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage.
Steering Committee of the Physicians' Health Study Research Group. (1988). Preliminary report: Findings from the aspirin component of the ongoing Physicians' Health Study. *New England Journal of Medicine,* **318**, 262–264.
Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *Journal of the American Statistical Association,* **54**, 30–34.
Wilkinson, L. (1988). *SYSTAT: The system for statistics.* Evanston, IL: SYSTAT.

Chronological Listing of Studies in Meta-Analysis

Carington, W. (1935). Preliminary experiments in precognitive guessing. *Journal of the Society for Psychical Research,* **29**, 86–104.
Rhine, J. B. (1938). Experiments bearing on the precognition hypothesis: I. Pre-shuffling card calling. *Journal of Parapsychology,* **2**, 38–54.
Rhine, J. B., Smith, B. M., & Woodruff, J. L. (1938). Experiments bearing on the precognition hypothesis: II. The role of ESP in the shuffling of cards. *Journal of Parapsychology,* **2**, 119–131.
Humphrey, B. M., & Pratt, J. G. (1941). A comparison of five ESP test procedures. *Journal of Parapsychology,* **5**, 267–293.
Rhine, J. B. (1941). Experiments bearing upon the precognition hypothesis: III. Mechanically selected cards. *Journal of Parapsychology,* **5**, 1–57.
Stuart, C. E. (1941). An analysis to determine a test predictive of extra-chance scoring in card-calling tests. *Journal of Parapsychology,* **5**, 99–137.
Humphrey, B. M., & Rhine, J. B. (1942). A confirmatory study of salience in precognition tests. *Journal of Parapsychology,* **6**, 190–219.

Rhine, J. B. (1942). Evidence of precognition in the covariation of salience ratios. *Journal of Parapsychology,* **6**, 111–143.
Nicol, J. F., & Carington, W. (1947). Some experiments in willed die-throwing. *Proceedings of the Society for Psychical Research,* **48**, 164–175.
Thouless, R. H. (1949). A comparative study of performance in three psi tasks. *Journal of Parapsychology,* **13**, 263–273.
Bastin, E. W., & Green, J. M. (1953). Some experiments in precognition. *Journal of Parapsychology,* **17**, 137–143.
McMahan, E. A., & Bates, E. K. (1954). Report of further Marchesi experiments. *Journal of Parapsychology,* **18**, 82–92.
Mangan, G. L. (1955). Evidence of displacement in a precognition test. *Journal of Parapsychology,* **19**, 35–44.
Osis, K. (1955). Precognition over time intervals of one to thirty-three days. *Journal of Parapsychology,* **19**, 82–91.
Nielsen, W. (1956). An exploratory precognition experiment. *Journal of Parapsychology,* **20**, 33–39.
Nielsen, W. (1956). Mental states associated with success in precognition. *Journal of Parapsychology,* **20**, 96–109.
Fahler, J. (1957). ESP card tests with and without hypnosis. *Journal of Parapsychology,* **21**, 179–185.
Mangan, G. L. (1957). An ESP experiment with dual-aspect targets involving one trial a day. *Journal of Parapsychology,* **21**, 273–283.
Anderson, M., & White, R. (1958). A survey of work on ESP and teacher-pupil attitudes. *Journal of Parapsychology,* **22**, 246–268.
Nash, C. B. (1958). Correlation between ESP and religious value. *Journal of Parapsychology,* **22**, 204–209.
Anderson, M. (1959). A precognition experiment comparing time intervals of a few days and one year. *Journal of Parapsychology,* **23**, 81–89.
Anderson, M., & Gregory, E. (1959). A two-year program of tests for clairvoyance and precognition with a class of public school pupils. *Journal of Parapsychology,* **23**, 149–177.
Nash, C. B. (1960). Can precognition occur diametrically? *Journal of Parapsychology,* **24**, 26–32.
Freeman, J. A. (1962). An experiment in precognition. *Journal of Parapsychology,* **26**, 123–130.
Rhine, J. B. (1962). The precognition of computer numbers in a public test. *Journal of Parapsychology,* **26**, 244–251.
Ryzl, M. (1962). Training the psi faculty by hypnosis. *Journal of the Society for Psychical Research,* **41**, 234–252.
Sanders, M. S. (1962). A comparison of verbal and written responses in a precognition experiment. *Journal of Parapsychology,* **26**, 23–34.
Freeman, J. (1963). Boy-girl differences in a group precognition test. *Journal of Parapsychology,* **27**, 175–181.
Rao, K. R. (1963). Studies in the preferential effect: II. A language ESP test involving precognition and "intervention." *Journal of Parapsychology,* **27**, 147–160.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4
Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

FREEMAN, J. (1964). A precognition test with a high-school science club. *Journal of Parapsychology*, **28**, 214–221.
FREEMAN, J., & NIELSEN, W. (1964). Precognition score deviations as related to anxiety levels. *Journal of Parapsychology*, **28**, 239–249.
SCHMEIDLER, G. (1964). An experiment on precognitive clairvoyance: Part I. The main results. *Journal of Parapsychology*, **28**, 1–14.
FREEMAN, J. A. (1965). Differential response of the sexes to contrasting arrangements of ESP target material. *Journal of Parapsychology*, **29**, 251–258.
OSIS, K., & FAHLER, J. (1965). Space and time variables in ESP. *Journal of the American Society for Psychical Research*, **59**, 130–145.
FAHLER, J., & OSIS, K. (1966). Checking for awareness of hits in a precognition experiment with hypnotized subjects. *Journal of the American Society for Psychical Research*, **60**, 340–346.
FREEMAN, J. A. (1966). Sex differences and target arrangement: High-school booklet tests of precognition. *Journal of Parapsychology*, **30**, 227–235.
ROGERS, D. P. (1966). Negative and positive affect and ESP run-score variance. *Journal of Parapsychology*, **30**, 151–159.
ROGERS, D. P., & CARPENTER, J. C. (1966). The decline of variance of ESP scores within a testing session. *Journal of Parapsychology*, **30**, 141–150.
BRIER, B. (1967). A correspondence ESP experiment with high-I.Q. subjects. *Journal of Parapsychology*, **31**, 143–148.
BUZBY, D. E. (1967). Subject attitude and score variance in ESP tests. *Journal of Parapsychology*, **31**, 43–50.
BUZBY, D. E. (1967). Precognition and a test of sensory perception. *Journal of Parapsychology*, **31**, 135–142.
FREEMAN, J. A. (1967). Sex differences, target arrangement, and primary mental abilities. *Journal of Parapsychology*, **31**, 271–279.
HONORTON, C. (1967). Creativity and precognition scoring level. *Journal of Parapsychology*, **31**, 29–42.
CARPENTER, J. C. (1968). Two related studies on mood and precognition run-score variance. *Journal of Parapsychology*, **32**, 75–89.
DUVAL, P., & MONTREDON, E. (1968). ESP experiments with mice. *Journal of Parapsychology*, **32**, 153–166.
FEATHER, S. R., & BRIER, R. (1968). The possible effect of the checker in precognition tests. *Journal of Parapsychology*, **32**, 167–175.
FREEMAN, J. A. (1968). Sex differences and primary mental abilities in a group precognition test. *Journal of Parapsychology*, **32**, 176–182.
NASH, C. S., & NASH, C. B. (1968). Effect of target selection, field dependence, and body concept on ESP performance. *Journal of Parapsychology*, **32**, 248–257.
RHINE, L. E. (1968). Note on an informal group test of ESP. *Journal of Parapsychology*, **32**, 47–53.
RYZL, M. (1968). Precognition scoring and attitude toward ESP. *Journal of Parapsychology*, **32**, 1–8.
RYZL, M. (1968). Precognition scoring and attitude. *Journal of Parapsychology*, **32**, 183–189.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

CARPENTER, J. C. (1969). Further study on a mood adjective check list and ESP run-score variance. *Journal of Parapsychology*, **33**, 48–56.
DUVAL, P., & MONTREDON, E. (1969). Precognition in mice: A confirmation. *Journal of Parapsychology*, **33**, 71–72.
FREEMAN, J. A. (1969). The psi-differential effect in a precognition test. *Journal of Parapsychology*, **33**, 206–212.
FREEMAN, J. A. (1969). A precognition experiment with science teachers. *Journal of Parapsychology*, **33**, 307–310.
JOHNSON, M. (1969). Attitude and target differences in a group precognition test. *Journal of Parapsychology*, **33**, 324–325.
MONTREDON, E., & ROBINSON, A. (1969). Further precognition work with mice. *Journal of Parapsychology*, **33**, 162–163.
SCHMIDT, H. (1969). Precognition of a quantum process. *Journal of Parapsychology*, **33**, 99–108.
BENDER, H. (1970). Differential scoring of an outstanding subject on GESP and clairvoyance. *Journal of Parapsychology*, **34**, 272–273.
FREEMAN, J. A. (1970). Sex differences in ESP response as shown by the Freeman picture-figure test. *Journal of Parapsychology*, **34**, 37–46.
FREEMAN, J. A. (1970). Ten-page booklet tests with elementary-school children. *Journal of Parapsychology*, **34**, 192–196.
FREEMAN, J. (1970). Shift in scoring direction with junior-high-school students: A summary. *Journal of Parapsychology*, **34**, 275.
FREEMAN, J. A. (1970). Mood, personality, and attitude in precognition tests. *Journal of Parapsychology*, **34**, 322.
HARALDSSON, E. (1970). Subject selection in a machine precognition test. *Journal of Parapsychology*, **34**, 182–191.
HARALDSSON, E. (1970). Precognition of a quantum process: A modified replication. *Journal of Parapsychology*, **34**, 329–330.
NIELSEN, W. (1970). Relationships between precognition scoring level and mood. *Journal of Parapsychology*, **34**, 93–116.
SCHMIDT, H. (1970). Precognition test with a high-school group. *Journal of Parapsychology*, **34**, 70.
BELOFF, J., & BATE, D. (1971). An attempt to replicate the Schmidt findings. *Journal of the Society for Psychical Research*, **46**, 21–31.
HONORTON, C. (1971). Automated forced-choice precognition tests with a "sensitive." *Journal of the American Society for Psychical Research*, **65**, 476–481.
MITCHELL, E. D. (1971). An ESP test from Apollo 14. *Journal of Parapsychology*, **35**, 89–107.
SCHMIDT, H., & PANTAS, L. (1971). Psi tests with psychologically equivalent conditions and internally different machines. *Journal of Parapsychology*, **35**, 326–327.
STANFORD, R. G. (1971). Extrasensory effects upon "memory." *Journal of the American Society for Psychical Research*, **64**, 161–186.
STEILBERG, B. J. (1971). Investigation of the paranormal gifts of the Dutch sensitive Lida T. *Journal of Parapsychology*, **35**, 219–225.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

THOULESS, R. H. (1971). Experiments on psi self-training with Dr. Schmidt's pre-cognitive apparatus. *Journal of the Society for Psychical Research*, **46**, 15–21.

HONORTON, C. (1972). Reported frequency of dream recall and ESP. *Journal of the American Society for Psychical Research*, **66**, 369–374.

JOHNSON, M., & NORDBECK, B. (1972). Variation in the scoring behavior of a "psychic" subject. *Journal of Parapsychology*, **36**, 122–132.

KELLY, E. F., & KANTHAMANI, B. K. (1972). A subject's efforts toward voluntary control. *Journal of Parapsychology*, **36**, 185–197.

SCHMIDT, H., & PANTAS, L. (1972). Psi tests with internally different machines. *Journal of Parapsychology*, **36**, 222–232.

CRAIG, J. G. (1973). The effect of contingency on precognition in the rat. *Research in Parapsychology 1972*, 154–156.

FREEMAN, J. A. (1973). The psi quiz: A new ESP test. *Research in Parapsychology 1972*, 132–134.

ARTLEY, B. (1974). Confirmation of the small-rodent precognition work. *Journal of Parapsychology*, **38**, 238–239.

HARRIS, S., & TERRY, J. (1974). Precognition in a water-deprived Wistar rat. *Journal of Parapsychology*, **38**, 239.

RANDALL, J. L. (1974). An extended series of ESP and PK tests with three English schoolboys. *Journal of the Society for Psychical Research*, **47**, 485–494.

EYSENCK, H. J. (1975). Precognition in rats. *Journal of Parapsychology*, **39**, 222–227.

HARALDSSON, E. (1975). Reported dream recall, precognitive dreams, and ESP. *Research in Parapsychology 1974*, 47–48.

HONORTON, C., RAMSEY, M., & CABIBBO, C. (1975). Experimenter effects in extrasensory perception. *Journal of the American Society for Psychical Research*, **69**, 135–149.

KANTHAMANI, H., & RAO, H. H. (1975). Response tendencies and stimulus structure. *Journal of Parapsychology*, **39**, 97–105.

LEVIN, J. A. (1975). A series of psi experiments with gerbils. *Journal of Parapsychology*, **39**, 363–365.

TERRY, J. C., & HARRIS, S. A. (1975). Precognition in water-deprived rats. *Research in Parapsychology 1974*, 81.

DAVIS, J. W., & HAIGHT, J. (1976). Psi experiments with rats. *Journal of Parapsychology*, **40**, 54–55.

JACOBS, J., & BREEDERVELD, H. (1976). Possible influences of birth order on ESP ability. *Research Letter* (Parapsychology Laboratory, University of Utrecht), No. 7, 10–20.

NEVILLE, R. C. (1976). Some aspects of precognition testing. *Research in Parapsychology 1975*, 29–31.

DRUCKER, S. A., DREWES, A. A., & RUBIN, L. (1977). ESP in relation to cognitive development and IQ in young children. *Journal of the American Society for Psychical Research*, **71**, 289–298.

HARALDSSON, E. (1977). ESP and the defense mechanism test (DMT): A further validation. *European Journal of Parapsychology*, **2**, 104–114.

SARGENT, C. L. (1977). An experiment involving a novel precognition task. *Journal of Parapsychology*, **41**, 275–293.

BIERMAN, D. J. (1978). Testing the "advanced wave" hypothesis: An attempted replication. *European Journal of Parapsychology*, **2**, 206–212.

BRAUD, W. (1979). Project Chicken Little: A precognition experiment involving the SKYLAB space station. *European Journal of Parapsychology*, **3**, 149–165.

HARALDSSON, E., & JOHNSON, M. (1979). ESP and the defense mechanism test (DMT) Icelandic study No. III: A case of the experimenter effect? *European Journal of Parapsychology*, **3**, 11–20.

O'BRIEN, J. T. (1979). An examination of the checker effect. *Research in Parapsychology 1978*, 153–155.

CLEMENS, D. B., & PHILLIPS, D. T. (1980). Further studies of precognition in mice. *Research in Parapsychology 1979*, 156.

HARALDSSON, E. (1980). Scoring in a precognition test as a function of the frequency of reading on psychical phenomena and belief in ESP. *Research Letter* (Parapsychology Laboratory, University of Utrecht), No. 10, 1–8.

SARGENT, C., & HARLEY, T. A. (1981). Three studies using a psi-predictive trait variable questionnaire. *Journal of Parapsychology*, **45**, 199–214.

WINKELMAN, M. (1981). The effect of formal education on extrasensory abilities: The Ozolco study. *Journal of Parapsychology*, **45**, 321–336.

NASH, C. B. (1982). ESP of present and future targets. *Journal of the Society for Psychical Research*, **51**, 374–377.

THALBOURNE, M., BELOFF, J., & DELANOY, D. (1982). A test for the "extraverted sheep versus introverted goats" hypothesis. *Research in Parapsychology 1981*, 155–156.

CRANDALL, J. E., & HITE, D. D. (1983). Psi-missing and displacement: Evidence for improperly focused psi? *Journal of the American Society for Psychical Research*, **77**, 209–228.

SCHWARTZ, S. A., & DEMATTEI, R. J. (1983). The Mobius Psi-Q test: Preliminary findings. *Research in Parapsychology 1982*, 103–105.

JOHNSON, M., & HARALDSSON, E. (1984). The Defense Mechanism Test as a predictor of ESP scores: Icelandic studies IV and V. *Journal of Parapsychology*, **48**, 185–200.

TEDDER, W. (1984). Computer-based long-distance ESP: An exploratory examination (RB/PS). *Research in Parapsychology 1983*, 100–101.

HESELTINE, G. L. (1985). PK success during structured and nonstructured RNG operation. *Journal of Parapsychology*, **49**, 155–163.

HARALDSSON, E., & JOHNSON, M. (1986). The Defense Mechanism Test (DMT) as a predictor of ESP performance: Icelandic studies VI and VII. *Research in Parapsychology 1985*, 43–44.

VASSY, Z. (1986). Experimental study of complexity dependence in precognition. *Journal of Parapsychology*, **50**, 235–270.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

HONORTON, C. (1987). Precognition and real-time ESP performance in a computer task with an exceptional subject. *Journal of Parapsychology*, **51**, 291–320.

Psychophysical Research Laboratories
P. O. Box 569
Plainsboro, NJ 08536

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

PSI COMMUNICATION IN THE GANZFELD

EXPERIMENTS WITH AN AUTOMATED TESTING SYSTEM AND A COMPARISON WITH A META-ANALYSIS OF EARLIER STUDIES

By Charles Honorton, Rick E. Berger, Mario P. Varvoglis, Marta Quant, Patricia Derr, Ephraim I. Schechter, and Diane C. Ferrari

ABSTRACT: A computer-controlled testing system was used in 11 experiments on ganzfeld psi communication. The automated ganzfeld system controls target selection and presentation, subjects' blind-judging, and data recording and storage. Videotaped targets included video segments (dynamic targets) as well as single images (static targets). Two hundred and forty-one volunteer subjects completed 355 psi ganzfeld sessions. The subjects, on a blind basis, correctly identified randomly selected and remotely viewed targets to a statistically significant degree, $z = 3.89$, $p = .00005$. Study outcomes were homogeneous across the 11 series and eight different experimenters. Performance on dynamic targets was highly significant, $z = 4.62$, $p = .0000019$, as was the difference between dynamic and static targets, $p = .002$. Suggestively stronger performance occurred with friends than with unacquainted sender/receiver pairs, $p = .0635$. The automated ganzfeld study outcomes are compared with a meta-analysis of 28 earlier ganzfeld studies. The two data sets are consistent on four dimensions: overall success rate, impact of dynamic and static targets, effect of sender/receiver acquaintance, and prior ganzfeld experience. The combined z for all 39 studies is 7.53, $p = 9 \times 10^{-14}$.

Research on psi communication in the ganzfeld developed as the result of earlier research suggesting that psi functioning is frequently associated with internal attention states brought about

This work was supported by the James S. McDonnell Foundation of St. Louis, Missouri, and by the John E. Fetzer Foundation of Kalamazoo, Michigan.

We wish to thank Marilyn J. Schlitz, Peter Rojcewicz, and Rosemarie Pilkington for their help in recruiting participants; Daryl J. Bem of Cornell University and Donald McCarthy of St. Johns University for helpful comments on an earlier draft of this paper; Edwin C. May of SRI International for performing the audio spectrum analysis; and Robert Rosenthal of Harvard University for suggestions concerning data analysis. We also wish to thank several PRL colleagues who contributed in various ways to the work reported here: Nancy Sondow for assistance in the preparation relaxation exercise and instruction tape that was used throughout, and George Hansen and Linda Moore who served frequently as lab senders. Hansen also provided technical assistance and conducted a data audit resulting in the correction of several minor errors that appeared in a version of this report presented at the 32nd Annual Convention of the Parapsychological Association. Finally, we thank the 241 volunteer participants for providing us with such interesting data.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

through dreaming, hypnosis, meditation, and similar naturally occurring or artificially induced states (Braud, 1978; Honorton, 1977). This generalization, based on converging evidence from spontaneous case studies, clinical observations, and experimental studies, led to the development of a low-level descriptive model of psi functioning, according to which, internal attention states facilitate psi detection by attenuating sensory and somatic stimuli that normally mask weaker psi input (Honorton, 1977, 1978). This "noise-reduction" model thus identified sensory deprivation as a key to the frequent association between psi communication and internal attention states, and the ganzfeld procedure was developed specifically to test the impact of perceptual isolation on psi performance.

Fifteen years have passed since the initial reports of psi communication in the ganzfeld (Braud, Wood, & Braud, 1975; Honorton & Harper, 1974; Parker, 1975). Dozens of additional psi ganzfeld studies have appeared since then, and the success of the paradigm has triggered substantial critical interest. Indeed, there is at least one critical review or commentary for every ganzfeld study reporting significant evidence of psi communication (Akers, 1984; Alcock, 1986; Blackmore, 1980, 1987; Child, 1986; Druckman & Swets, 1988; Harley & Matthews, 1987; Harris & Rosenthal, 1988; Honorton, 1979, 1983, 1985; Hövelmann, 1986; Hyman, 1983, 1985, 1988; Hyman & Honorton, 1986; Kennedy, 1979; McClenon, 1986; Palmer, 1986; Palmer, Honorton, & Utts, 1989; Parker & Wiklund, 1987; Rosenthal, 1986; Sargent, 1987; Scott, 1986; Stanford, 1984, 1986; Stokes, 1986; Utts, 1986).

Of the many controversies spanning the history of parapsychological inquiry, the psi ganzfeld domain is unique in three respects. First, the central issue involves the replicability of a theoretically based technique rather than the special abilities of exceptional individuals (Honorton, 1977). Second, meta-analytic techniques have been used to assess statistical significance, effect size, and potential threats to validity (Harris & Rosenthal, 1988; Honorton, 1985; Hyman, 1985, 1988; Rosenthal, 1986). Third, investigators and critics have agreed on specific guidelines for the conduct and evaluation of future psi ganzfeld research (Hyman & Honorton, 1986).

The Automated Ganzfeld Testing System

Psi ganzfeld experiments typically involve four participants. The subject (or receiver, R) attempts to gain target-relevant imagery while in the ganzfeld; following the ganzfeld/imagery period, R

tries—on a blind basis—to identify the actual target from among four possibilities. A physically isolated sender (Se) views the target and attempts to communicate salient aspects of it to R. Two experimenters (Es) are usually required. One E manages R, elicits R's verbal report of ganzfeld imagery (mentation), and supervises R's blind judging of the target and decoys; a second E supervises Se, and randomly selects and records the target.

We developed an automated ganzfeld testing system ("autoganzfeld") to eliminate potential methodological problems that were identified in earlier ganzfeld studies (Honorton, 1979; Hyman & Honorton, 1986; Kennedy, 1979) and to explore factors associated with successful performance. The system provides computer control of target selection and presentation, blind judging, subject feedback, and data recording and storage (Berger & Honorton, 1986). A computer-controlled videocassette recorder (VCR) accesses and automatically presents target stimuli to Se. A second E is required only for assistance in target selection The system includes an experimental design module through which E specifies the sample size and status of a new series.

The system was designed to enable further assessment of factors identified with successful performance in earlier ganzfeld studies. Differences in target type and sender/receiver acquaintance seem to be particularly important. Significantly better performance occurred in studies using dynamic rather than static targets. Dynamic targets contain multiple images reinforcing a central theme, whereas static targets contain a single image. Also, studies permitting subjects to have friends as their senders yielded significantly superior performance compared to those requiring subjects to work with laboratory senders. (See "Comparison of Study Outcomes with Ganzfeld Meta-Analysis" in the Results section.)

The autoganzfeld system uses both dynamic and static targets. The dynamic targets are excerpts from films; static targets include art work and photographs. Receivers may, if they choose, bring friends or family members to serve as their senders; a session setup module registers the sender type and other session information.

In this report, we present the results of the 11 autoganzfeld series conducted between the inauguration of the experiments in February, 1983, and September, 1989, when funding problems required suspension of the PRL research program.[1] We focus on

[1] This article conforms to the reporting guidelines recommended by Hyman and Honorton (1986). Because of the size of this database, however, it is not practical to

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

(1) evidence for psi in the autoganzfeld situation, (2) the impact of dynamic versus static targets, (3) the effects of sender/receiver acquaintance, (4) the impact of prior psi ganzfeld experience, and (5) a comparison of these four factors with the outcomes of earlier nonautomated psi ganzfeld experiments. Our findings on demographic, psychological, and target factors will be presented in later reports.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Subjects

The participants are 100 men and 141 women ranging in age from 17 to 74 years (mean = 37.3, SD = 11.8). This is a well-educated group; the mean formal education is 15.6 years (SD = 2.0).

Our primary sources of recruitment include referrals from colleagues (24%), media presentations concerning PRL research (23%), friends or acquaintances of PRL staff (20%), and referrals from other participants (18%).

Belief in psi is strong in this population. On a seven-point scale where "1" indicates strong disbelief and "7" indicates strong belief in psi, the mean is 6.20 (SD = 1.03); only two participants rated their belief in psi below the midpoint of the scale. Personal experiences suggestive of psi were reported by 88% of the subjects; 80% reported ostensible telepathic experiences. Eighty percent of the participants have had some training in meditation or other techniques involving internal focus of attention.

Participant Orientation

Initial contact. New participants receive an information pack before their first session. The information pack includes a 55-item personal history survey (Participant Information Form [PIF]; Psychophysical Research Laboratories, 1983), Form F of the Myers-Briggs Type Indicator (MBTI; Briggs & Myers, 1957), general information about the research program, and directions for reaching PRL. Participants usually return the completed questionnaires before their first session. However, if new participants are scheduled on short notice, they either complete the questionnaires at PRL or, in a few cases, at home after the session.

include the data in an appendix to the report. Instead, we will supply the data to qualified investigators in a Lotus-compatible, MS-DOS computer disk file. There is a small fee to cover materials and mailing. Address inquiries to the *Journal.*

Whenever possible, new participants are encouraged to come in for a preliminary orientation session, prior to their first PRL ganzfeld session. The orientation serves as a "get acquainted" session for participants and the PRL staff, and introduces participants to the PRL program and facility. Participants who avail themselves of this option generally complete the MBTI and PIF questionnaires during the orientation session. We inform new participants that they may bring a friend or family member to serve as their sender. When a participant chooses not to do so, a PRL staff member serves as sender. We encourage participants to reschedule their session rather than feel they must come in to "fulfill an obligation" if they are not feeling well.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Session orientation. We greet participants at the door when they arrive and attempt to create a friendly and informal social atmosphere. Coffee, tea, and soft drinks are available. E and other staff members engage in conversation with R during this period. When a laboratory sender is used, time is taken for sender and receiver to become acquainted.

If the participant is a novice, we describe the rationale and background of the ganzfeld research, and we seek to create positive expectations concerning R's ability to identify the target. This information is tailored to our perception of the needs of the individual participant, but it generally includes four elements: (1) a brief review of experimental, clinical, and spontaneous case trends indicating that ESP is more readily detected during internal attention states such as dreaming, hypnosis, and meditation (Honorton, 1977), (2) the notion that these states all involve physical relaxation and functional sensory deprivation, suggesting that weak ESP impressions may be more readily detected when perceptual and somatic noise is reduced, (3) the development of the ganzfeld technique to test this noise-reduction hypothesis, and (4) the long-term success of the ganzfeld technique as a means of facilitating psi communication in unselected subjects.

We encourage "goal orientation" and discourage excessive "task orientation" during the session; this is especially emphasized with participants who appear to be anxious or overly concerned about their ability to succeed in the ganzfeld task. We discourage participants from analyzing their mentation during the session, and tell them that they will have an opportunity to analyze their mentation during the judging procedure. They are encouraged to adopt the role of an outside observer of their mental processes during the ganzfeld. Again, this is emphasized with those who appear anxious

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

about their performance; they are advised to relax, follow the taped instructions, and to simply allow the procedure to work. We inform participants that they may experience various types of correspondence between their mentation and the target; they are told that they may experience direct, literal correspondences to the target, but that they should also be prepared for correspondences involving distortions or transformations of the target content, cognitive associations, and similarities in emotional tone. Finally, we orient new participants to where Se and E will be located during the session.

Method

Layout and Equipment

R and Se are sequestered in nonadjacent, sound-isolated and electrically shielded rooms. Both rooms are copper-screened, and are 14 ft apart on opposite sides of E's monitoring room, which provides the only access. R and Se remain isolated in their respective rooms until R completes the blind-judging procedure.

R's room is an Industrial Acoustics Corp., IAC 1205A Sound-Isolation Room, consisting of two 4-inch sheetrock-filled steel panels. The two panels are separated by a 4-inch air space, for a total thickness of one foot.

The inside walls and ceiling of Se's room are covered with 4-inch Sonex® acoustical material, similar to that used in commercial broadcast studios. A free-standing Sonex-covered plywood barrier (5 ft wide by 8 ft high) positioned inside the sender's room, between Se's chair and the acoustical door, blocks sound transmission through the door frame. Figure 1 shows the floor plan of the experimental rooms.

E occupies a console housing the computer system and other equipment. The computer is an Apple II Plus with two disk drives, a printer, and an expansion chassis. The computer peripherals include a real-time clock, a noise-based random number generator (RNG), a Cavri Interactive Video Interface™, an Apple game paddle, and a fan. Other equipment includes a color TV monitor, the VCR used to access and display targets, and three electrically isolated audiocassette recorders. One audiocassette recorder presents audio stimuli (prerecorded relaxation exercises, session instructions, and white noise). Another plays background music during the experimental setup. The third records R's ganzfeld mentation and judging period associations. There is two-way intercom communication between E and R. One-way audio communication from R to Se allows Se to listen to R's ganzfeld mentation.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

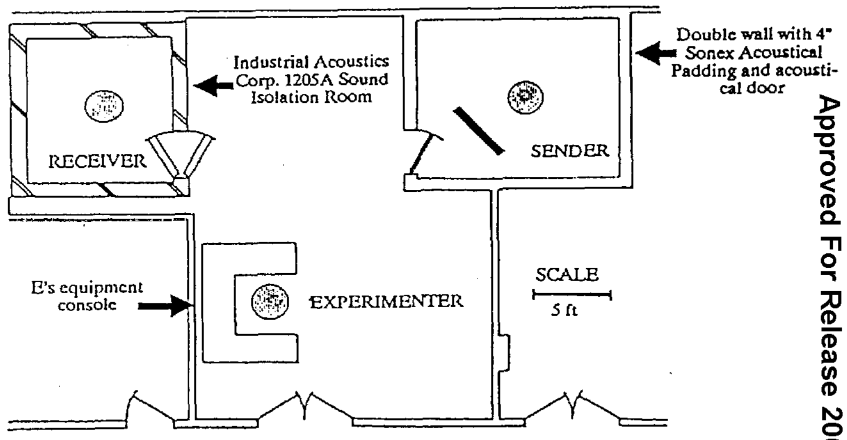


Figure 1. Floor plan of experimental suite.

Receiver Preparation

R sits in a comfortable reclining chair in the IAC room. Se keeps R company while E prepares R for visual and auditory ganzfeld stimulation. Translucent hemispheres are taped over R's eyes with Micropore® tape. Headphones are placed over R's ears. A clip-on microphone is fastened to R's collar. A 600-watt red-filtered floodlight, located approximately 6 ft in front of R's face, is adjusted in intensity until R reports a comfortable, shadow-free, homogeneous visual field. White noise level is similarly adjusted; R is informed that the white noise should be as loud as possible without being annoying or uncomfortable. The ganzfeld light and white noise intensity are adjusted from E's console after R and Se are sequestered in their respective rooms.

Sender Preparation

Se sits in a comfortable reclining chair in the sender's room. Se faces a color TV monitor, wearing headphones. During the session, Se can hear R's mentation report through one headphone; if dynamic targets are used, Se hears the target audio channel through the other headphone.

Series Manager Setup Procedures

E accesses the autoganzfeld computer program through the *Series Manager* software. *Series Manager* is a password-protected, menu-driven control program. It provides the only means through which an experimenter may specify parameters for the series design, register new participants in the series, set up a session, and run a session. The *Series Manager* menu is accessed through entry of a private (and nonechoing) password.

Series design. A valid series design must exist before sessions can be run in an experimental series. This is done through the *Series Manager* "design" module. The design module prompts E to specify the type of series (pilot, screening, or formal), the number of participants, the maximum number of trials per participant, the total number of trials per series, and the series name. There is no provision for changing the series design once it is accepted by E. Design parameters are saved in a disk file; they are passed to the experimental program at the beginning of the session.

Participant registration. When R is new to a series, E accesses "Participant Registration" from the *Series Manager* menu before the session. E is prompted to enter R's name and identification number. The module verifies that the maximum number of participants specified in the design is not exceeded. (An error message appears if an attempt is made to register more participants than are specified in the design; then, control is returned to the *Series Manager* menu.)

Session setup. E then selects "Session Setup" from the *Series Manager* menu. E is prompted to enter R's name and the program verifies that R has not already completed the maximum number of trials specified in the design module. (An error message appears if a participant has completed the number of sessions allowed for the series or has not been properly registered; control is then returned to the *Series Manager* menu.) E enters Se's name and the sender type: lab, lab friend, or friend. *Lab senders* are PRL staff members whose acquaintance with the participant is limited to the experiment. *Lab friend* refers to PRL staff senders who have some social acquaintance with R outside the laboratory. *Friend senders* are friends or family members of the participant. Finally, E enters the ganzfeld light and noise intensity levels and his or her initials. E then leaves

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the monitoring room while another PRL staff person supervises target selection.

Targets

The system uses short video segments (*dynamic targets*) and still pictures (*static targets*) as targets. Dynamic targets include excerpts from motion pictures, documentaries, and cartoons. Static targets include art prints, photographs, and magazine advertisements.

There are 160 targets, arranged in judging sets of four dynamic or four static targets. The sets were constructed to minimize similarities among targets within a set. The targets are recorded on four one-half-inch VHS format videocassettes; each videocassette contains 10 target sets (5 dynamic and 5 static). A signal recorded on an audio track of each videocassette allows computer access of the targets. Target display time—to Se during each sending period and to R during the judging period—is approximately one minute; blank space added to briefer targets insures that the VCR remains in play mode for the same length of time for all targets.

Preview packs. The video display format of the autoganzfeld targets does not permit simultaneous viewing of the entire target set during the judging procedure as is done in many nonautomated ganzfeld studies. Each target set is therefore accompanied by a preview pack containing brief excerpts of all four targets in the set; this gives R a general impression of the range of target possibilities. R views the preview pack at the beginning of the judging procedure; it runs approximately 30 sec.

Target Selection

The target selector (TS) is a PRL staff member who has no contact with either E or R until after the blind-judging procedure. TS is needed to load the videocassette containing the target into the VCR. TS is informed which of the four videocassettes contains the target, but remains blind to the target's identity. If Se is a staff member, Se serves this role; otherwise, a staff member not involved in the session serves as TS. (In the latter case, Se and R are sequestered in their respective rooms before TS enters the monitoring room.)

The *Series Manager* program prompts TS to press a key on the computer keyboard. A program call to the hardware RNG obtains the target value (a number between 1 and 160) and stores it in com-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

puter memory.[2] The program determines the target set and videocassette number from the target value. The videocassette number is displayed on the monitor, and TS is prompted to insert it into the VCR. The program verifies that the correct videocassette has been inserted and clears the monitor screen; if the videocassette is not correct, an error message prompts TS to insert the correct videocassette.

TS places a cardboard cover over the VCR's front panel to conceal the digital counters and VU meters. Finally, TS leaves the monitoring room with the three remaining videocassettes, knocking three times on the monitoring room door as a signal for E to return.

Relaxation Exercises and Ganzfeld Instructions

R and Se undergo a 14-min prerecorded relaxation exercise before the mentation/sending period. This provides a unique shared experience for R and Se before the ESP task. The relaxation exercise includes progressive relaxation exercises and autogenic phrases (Jacobson, 1929; Shultz, 1950). Ganzfeld instructions are recorded after the relaxation exercise. The instructions and relaxation exercise are delivered in a slow, soothing but confident manner with ocean sounds in the background. The style of presentation is similar to a hypnotic induction procedure. The ganzfeld instructions to R, which are also heard by Se, are as follows:

> During this experiment we want you to think out loud. Report all of the images, thoughts, and feelings that pass through your mind. Do not cling to any of them. Just observe them as they go by. At some point during the session, we will send you the target information. Do not try to anticipate or conjure up this information. Just give yourself the suggestion, right now—in the form of making a wish—that the information will appear in consciousness at the appropriate time. Keep your eyes open as much as possible during the session and allow your consciousness to flow through the sound you will hear through the headphones. One of us will be monitoring you in the other room. Now get as comfortable as possible, release all conscious hold of your body, and allow it to relax completely. As soon as you begin observing your mental processes, start thinking out loud. Continue to share your thoughts, images, and feelings with us throughout the session.

[2] An exception occurs in the two target comparison series (Series 301 and 302). See pp. 112–113.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Mentation/Sending Procedures

Receiver mentation report. After the relaxation exercise and instructions, R listens to the white noise through headphones for 30 minutes. R reports whatever thoughts, images, and feelings occur in the ganzfeld. The mentation report is monitored by E and Se from their respective rooms. The mentation report is tape recorded, and E takes detailed notes for review from R prior to judging.

Target presentation and sender procedures. A Cavri Video Interface automates computer access and control of targets from a JVC BR-6400U VCR. An electronic video switcher selectively routes the video output (VCR or computer text mode) to three color TV monitors, one each for E, R, and Se. E's and R's monitors remain in computer text mode until the judging period. During each of the six sending periods, Se's TV monitor is switched from computer text to VCR mode.

At the beginning of each sending period, Se's monitor displays the prompt, "Silently communicate the contents and meaning of the target to [R's first name]." Se views the target and attempts to communicate its contents to R. Se mentally reinforces R for target-related associations and mentally discourages R when the mentation is unrelated to the target.

Judging Procedure

After the mentation period, E turns off the ganzfeld light and reads back R's mentation from the session notes. R remains in ganzfeld during the mentation review to minimize any abrupt shift in state. E's and R's TV monitors are switched into VCR mode by the computer, which also prompts Se to "Silently direct [R's first name] to select the target that you saw." Se's TV monitor remains blank (computer mode) during this period.

R removes the eye covers and views the preview pack. From their respective rooms, R and E then view the four potential targets (the actual target and three decoys), which are presented in one of four random sequences. R, viewing each candidate, associates to the item as though it were the actual target, describing perceived similarities between the item and the ganzfeld mentation. While R associates to each candidate, E points out potential correspondences that R may have overlooked.[3] R views any of the target candidates as often as desired before proceeding to the judging task.

[3] This applies to Pilot Series 3, Novice Series 103–105, and to Experienced Series

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

A 40-point rating scale then appears on R's TV monitor. The scale is labelled 0% on the left and 100% on the right. Using a computer-game paddle to move a pointer horizontally across the rating scale, R indicates the degree of similarity between his ganzfeld mentation and each potential target. E and Se view R's ratings on their monitors. The program checks for ties, and, if they occur, R re-rates the four candidates to obtain unique ratings for each. The program then converts R's ratings into ranks. A rank of 1 is assigned to the candidate R believes has the strongest similarity to his ganzfeld mentation; a rank of 4 is given to the candidate R believes is least like his ganzfeld experience.

Feedback and Post-Session Procedures

After R finishes judging, Se leaves the sender's room and enters R's room with E. Se reveals the actual target, which the computer automatically displays on R's TV monitor. The session data are written to a floppy disk file.

Following feedback, E is prompted to backup the series data disk. The target videocassette is then automatically wound to a position near the center of the videocassette (frame 50,000). E selects "Analysis" from the *Series Manager* menu and obtains a hardcopy printout of the session data file. The printout includes: the file name, R's name and ID number, series type, session number, Se's name, E's initials, date and start time, target number, target position in the set, R's target ranking, the standardized target rating (z score), target judging sequence, target name, target type and set number, sender type, light and white noise levels, finish time, and optional experimenter's comments. The printout is attached to E's notes on R's mentation and placed in a ring binder containing all such information for the series. The audio tape of the session is similarly filed.

Experimenters

Eight Es contributed to the autoganzfeld database. Honorton, one of the originators of the psi ganzfeld technique, has conducted psi ganzfeld experiments over a 16-year period. Derr and Varvoglis

201 and 302. It does not apply to the earlier series (Pilot Series 1–2; Novice Series 101–102; or Experienced Series 301). This practice was initiated because participants frequently failed to identify obvious correspondences between their mentation and target elements.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

worked with Honorton at Maimonides Medical Center and were trained by him. Berger is primarily responsible for the technical implementation of the autoganzfeld system. He trained Honorton, Derr, Varvoglis, and Schechter in its use. Honorton trained Quant, Ferrari, and Schlitz in the use of the autoganzfeld system.[4]

Experimental Series

Altogether, 241 participants contributed 355 sessions in 11 series. To fully address the issue of selective reporting, we include every session completed from the inauguration of the experiments in February, 1983, to September, 1989, when the PRL facility was closed. Thus, this database has no "file-drawer" problem (Rosenthal, 1984).

The studies include three pilot series and eight formal series. Five of the formal series were single-session studies with novice participants. The remaining three formal series involved experienced participants.

Pilot Series

Series 1. This initial pilot series was conducted during the development and testing of the autoganzfeld system. It served to test system operation, to detect and correct programming errors, and to fine-tune session timing functions. Nineteen subjects contributed 22 sessions as Rs. Seven, including PRL staff members, had prior experience as Rs in nonautomated ganzfeld studies at Maimonides Medical Center. The remaining 12 Rs were novices with no prior ganzfeld experience. Series sample size was not specified in advance; the series continued until we were satisfied that the system was operating reliably.

Series 2. This pilot series was designed by Berger in an attempt to avert potential displacement effects and subject judging problems by having E rather than R serve as judge. R received feedback only to the actual target. Four participants contributed to this series. Nine of the planned 50 sessions were completed before Berger's departure from PRL when this series was discontinued.

[4] Berger, Schechter, and Varvoglis have doctorate degrees in psychology. Quant holds a masters degree in counselling psychology, and Ferrari has a bachelors degree in psychology. Schlitz has conducted independent ganzfeld and remote-viewing research in other laboratories and has a masters degree in anthropology.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

A 40-point rating scale then appears on R's TV monitor. The scale is labelled 0% on the left and 100% on the right. Using a computer-game paddle to move a pointer horizontally across the rating scale, R indicates the degree of similarity between his ganzfeld mentation and each potential target. E and Se view R's ratings on their monitors. The program checks for ties, and, if they occur, R re-rates the four candidates to obtain unique ratings for each. The program then converts R's ratings into ranks. A rank of 1 is assigned to the candidate R believes has the strongest similarity to his ganzfeld mentation; a rank of 4 is given to the candidate R believes is least like his ganzfeld experience.

Feedback and Post-Session Procedures

After R finishes judging, Se leaves the sender's room and enters R's room with E. Se reveals the actual target, which the computer automatically displays on R's TV monitor. The session data are written to a floppy disk file.

Following feedback, E is prompted to backup the series data disk. The target videocassette is then automatically wound to a position near the center of the videocassette (frame 50,000). E selects "Analysis" from the *Series Manager* menu and obtains a hardcopy printout of the session data file. The printout includes: the file name, R's name and ID number, series type, session number, Se's name, E's initials, date and start time, target number, target position in the set, R's target ranking, the standardized target rating (z score), target judging sequence, target name, target type and set number, sender type, light and white noise levels, finish time, and optional experimenter's comments. The printout is attached to E's notes on R's mentation and placed in a ring binder containing all such information for the series. The audio tape of the session is similarly filed.

Experimenters

Eight Es contributed to the autoganzfeld database. Honorton, one of the originators of the psi ganzfeld technique, has conducted psi ganzfeld experiments over a 16-year period. Derr and Varvoglis

201 and 302. It does not apply to the earlier series (Pilot Series 1–2; Novice Series 101–102; or Experienced Series 301). This practice was initiated because participants frequently failed to identify obvious correspondences between their mentation and target elements.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

worked with Honorton at Maimonides Medical Center and were trained by him. Berger is primarily responsible for the technical implementation of the autoganzfeld system. He trained Honorton, Derr, Varvoglis, and Schechter in its use. Honorton trained Quant, Ferrari, and Schlitz in the use of the autoganzfeld system.[4]

Experimental Series

Altogether, 241 participants contributed 355 sessions in 11 series. To fully address the issue of selective reporting, we include every session completed from the inauguration of the experiments in February, 1983, to September, 1989, when the PRL facility was closed. Thus, this database has no "file-drawer" problem (Rosenthal, 1984).

The studies include three pilot series and eight formal series. Five of the formal series were single-session studies with novice participants. The remaining three formal series involved experienced participants.

Pilot Series

Series 1. This initial pilot series was conducted during the development and testing of the autoganzfeld system. It served to test system operation, to detect and correct programming errors, and to fine-tune session timing functions. Nineteen subjects contributed 22 sessions as Rs. Seven, including PRL staff members, had prior experience as Rs in nonautomated ganzfeld studies at Maimonides Medical Center. The remaining 12 Rs were novices with no prior ganzfeld experience. Series sample size was not specified in advance; the series continued until we were satisfied that the system was operating reliably.

Series 2. This pilot series was designed by Berger in an attempt to avert potential displacement effects and subject judging problems by having E rather than R serve as judge. R received feedback only to the actual target. Four participants contributed to this series. Nine of the planned 50 sessions were completed before Berger's departure from PRL when this series was discontinued.

[4] Berger, Schechter, and Varvoglis have doctorate degrees in psychology. Quant holds a masters degree in counselling psychology, and Ferrari has a bachelors degree in psychology. Schlitz has conducted independent ganzfeld and remote-viewing research in other laboratories and has a masters degree in anthropology.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Series 3. This pilot series was a practice series for participants who completed the allotted number of sessions in ongoing formal series but who wanted additional ganzfeld experience. This series also includes several demonstration sessions when TV film crews were present and provided receiver experience for new PRL staff. The sample size was not preset.

Novice ("First-Timers") Series

The identification of characteristics associated with successful initial performance was a major goal of the PRL ganzfeld project (Honorton & Schechter, 1987). Except for Series 105, each novice series includes 50 ganzfeld novices, that is, participants with no prior ganzfeld experience. Each novice contributed a single ganzfeld session. Most novices had not participated in *any* psi experiment prior to the novice series.

Series 101. This is the first novice series.

Series 102. Beginning with this series, R was prompted after the mentation period to estimate the number of minutes since the end of the relaxation/instructions tape.

Series 103. Starting with this series, Rs were given the option of having no sender (i.e., "clairvoyance" condition). Only four participants opted to have no sender.

Series 104. A visiting scientist (Marilyn Schlitz) served as E in seven sessions and as Se in six sessions with subjects from The Juilliard School in New York.

Series 105. This series was started to accommodate the overflow of Juilliard students from Series 104. The sample size was set to 25. Six sessions were completed at the time the PRL program was suspended. (There were 20 Juilliard students altogether. Sixteen were in Series 104 and four were in Series 105.)

Experienced Subjects Series

Series 201. This series involved especially promising subjects. The number of trials was set to 20. Seven sessions by three Rs were completed at the time the PRL program was suspended.

Series 301. This series compared dynamic and static targets. Sample size was set to 50 sessions. Twenty-five experienced subjects each contributed two sessions. The autoganzfeld program was modified for this series so that each R would have one session with dynamic targets and one session with static targets. Subjects were informed of this only after completing both sessions.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Series 302. This series used a single dynamic target set (Set 20). In earlier series, Target 77 ("Tidal Wave Engulfing Ancient City") had an especially strong success rate while Target 79 ("High-Speed Sex Trio") had never been correctly identified. We made two program modifications for this series. The target selection ("Randomize") routine was modified to select only targets in Set 20, and the VCR tape-centering routine was modified to wind the videotape to a randomly selected position between frame numbers 85,000 and 95,000. The second modification insured that E could not be cued, perhaps unconsciously, by the time required to wind the tape from its initial position to the target location.

The study involved experienced Rs who had no prior experience with Set 20. Each R contributed one session. Participants were unaware of the purpose of the study or that it was limited to one target set. The design called for the series to continue until 15 sessions were completed with each of the two targets of interest. Twenty-five sessions were completed when the PRL program was suspended.

Statistical Analysis

Except for two pilot series, series sample sizes were specified in advance. Our primary hypothesis was that the observed success rate—the proportion of correctly identified targets—would reliably exceed the null hypothesis expectation of .25. To test this hypothesis, we calculated the exact binomial probability for the observed number of direct hits (ranks of 1) with $p = .25$ and $q = .75$. On the basis of the overwhelmingly positive outcomes of earlier studies, we preset alpha to .05, one-tailed.

We also tested two secondary hypotheses, based on patterns of success in earlier psi ganzfeld research. These are: (1) that dynamic targets are significantly superior to static targets, and (2) that performance is significantly enhanced when the sender is a friend of R, compared to when R and Se are not acquainted. We initially planned to test these hypotheses by chi-square tests, a trial-based analysis. However, a consultant (Dr. Robert Rosenthal) suggested that a t test using the series as the unit would be a more powerful test of these hypotheses, and we have followed his recommendation. The remaining analyses are exploratory.[5]

[5] The statistical analyses in this report were performed using SYSTAT (Wilkin-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 1
OUTCOME BY SERIES

Series	Series type	N subjects	N trials	Hits N	Hits %	Effect size (h)	Effect size z
1	Pilot	19	22	8	36	.25	.99
2	Pilot	4	9	3	33	.18	.25
3	Pilot	25	36	10	28	.07	.22
101	Novice	50	50	12	24	−.02	−.30
102	Novice	50	50	18	36	.24	1.60
103	Novice	50	50	15	30	.11	.67
104	Novice	50	50	18	36	.24	1.60
105	Novice	6	6	4	67	.87	1.78
201	Experienced	3	7	3	43	.38	.69
301	Experienced	25	50	15	30	.11	.67
302	Experienced	25	25	16	64	.81	3.93
	Overall	241	355	122	34	.20	3.89

Note. The z scores are based on the exact binomial probability with $p = .25$ and $q = .75$.

RESULTS

Overall Success Rate

Ganzfeld hit rate. There were 241 participants, who contributed 355 autoganzfeld sessions. The 122 direct hits (34.4%) yield an exact binomial p of .00005 ($z = 3.89$). The effect size, Cohen's h (Cohen, 1977), is .20. The 95% confidence interval (CI) is a hit rate from 30% to 39%. Because this level of accuracy would occur about one time in 20,000 by chance, we reject the null hypothesis. (See Table 1.)

Success rate by series. Of the 11 series, 10 yield positive outcomes. The mean series effect size is .29, $SD = .29$, t (10) = 3.32.

Homogeneity of effect sizes. Traditionally, psi investigators have been preoccupied by whether there is a significant nonzero effect. An equally important issue, however, is the size of the effect. There is a growing tendency among behavioral scientists to define replicability in terms of the homogeneity of effect sizes (Hedges, 1987; Rosenthal, 1986; Utts, 1986). Two or more studies are replicates of one another if their effect sizes are homogeneous. We assess the homogeneity of effect sizes across the 11 series by performing a chi-square homogeneity test comparing the effect size for each series with the weighted mean effect size (Hedges, 1981; Rosenthal, 1984). The formula is:

$$\chi^2(k - 1) = \sum_{i=1}^{k} N_i(h_i - \bar{h})^2,$$

where k is the number of studies, N_i is the sample size of the ith study, and the weighted mean effect size is:

$$\bar{h} = \frac{\sum_{i=1}^{k} N_i h_i}{\sum_{i=1}^{k} N_i}.$$

The test shows that the series effect sizes are not significantly nonhomogeneous: $\chi^2 = 16.25$, 10 df, $p = .093$.

son, 1988). When t tests are reported on samples with unequal variances, they are calculated using the separate variances within groups for the error and degrees of freedom following Brownlee (1965). Combined zs are based on Stouffer's method (Rosenthal, 1984). Unless otherwise specified, p levels are one-tailed.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 2
OUTCOME BY EXPERIMENTER

Experimenter	N trials	Hits N	Hits %	Effect size (h)
Quant	106	38	36	.24
Honorton	72	27	38	.29
Berger	53	18	34	.20
Derr	45	12	27	.05
Varvoglis	43	11	26	.03
Schechter	14	5	36	.23
Ferrari	15	9	60	.72
Schlitz	7	2	29	.08

Homogeneity of Outcome by Experimenter

Eight Es contributed to the autoganzfeld database. (See Table 2.) All eight experimenters have positive effect sizes. A chi-square homogeneity test, using the mean effect sizes for each E weighted by sample size, indicates that the results are homogeneous across experimenters: $\chi^2 = 7.13$, 7 df, $p = .415$.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 3
GANZFELD SUCCESS IN RELATION TO NUMBER OF SESSIONS

	No. of sessions as receiver			
	1	2	3	4 +
N subjects	183	23	24	11
N trials	183	46	72	54
Hits	53	19	31	19
% Hits	29	41	43	35
Effect size (*h*)	.09	.34	.38	.22

Subject-Based Analysis

Seventy-six percent of the participants ($N = 183$) contributed a single session as R. Fifty-eight Rs contributed multiple sessions. Participants with multiple sessions either had direct hits or strongly suggestive target mentation correspondences in their first session. (See Table 3.)

Success rate by subjects. To test the consistency of ganzfeld performance across participants, we use the standardized ratings of the target and decoys (Stanford's *z* scores; Stanford & Sargent, 1983) as the dependent variable. Stanford *z*s are averaged for participants with multiple sessions. Direct hits and Stanford *z*s are highly correlated. In this database, N (353) is .776. The mean Stanford *z* for the 241 participants is .21 ($SD = 1.04$), and $t(240) = 3.22$ ($p = .00073$). The 95% CI is a Stanford *z* from .08 to .35. The effect size (Cohen's *d*; Cohen, 1977) is .21. (The effect size for subjects is nearly identical to the trial-based effect size, $h = .20$.) Thus, there is a general tendency for participants to give higher ratings to the actual target than to the decoys, and the significance of these experiments is not attributable to exceptional performance by a few outstanding subjects.

Dynamic Versus Static Targets

The success rate for dynamic targets is highly significant. There are 190 dynamic target sessions and 77 direct hits (40%, $h = .32$; exact binomial $p = 1.9 \times 10^{-6}$, $z = 4.62$). The hit rate for static targets is not significant (165 trials, 45 hits, 27%, $h = .05$, $p = .276$, $z = .59$). Using the series effect size as the outcome variable and target type as the predictor variable, the point-biserial correlation (r_p) between ganzfeld performance and target type is .663, $t(17) = 3.65$, $p = .002$.[6] The 95% CI for dynamic targets is a hit rate from 34% to 47%. The CI for static targets is from 21% to 34%. Thus, our hypothesis concerning the superiority of dynamic targets is strongly supported.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 4
SENDER/RECEIVER PAIRING

	Sender as:		
	Lab	Lab friend	Friend
N trials	140	66	145
N hits	46	24	52
% Hits	33	36	36
Effect size (*h*)	.18	.24	.24
z	2.01	1.93	2.83
p	.023	.026	.0023

Sender/Receiver Pairing

Receivers are more successful with friends than with laboratory senders, although the difference is not statistically significant. The number of sessions in this analysis is 351 because four subjects opted to have no sender. The best performance occurs with friend senders. Sessions with laboratory senders, although significant, have the lowest success rate. (See Table 4.)

Using series effect sizes as the unit of analysis and sender type as the predictor variable (combining lab friend and friends), r_p is .363, $t(17) = 1.61$, $p = .0635$.[7] The 95% CI for sessions with friends is a hit rate from 33.3% to 47%. For lab senders, the CI is from 18.3% to 41.8%. Thus, although the effect of sender type is not statistically significant, there is a trend toward better results with friends.

---

[6] Separate effect sizes were obtained for the dynamic and static target sessions of each series. Since Series 302 used dynamic targets only, the analysis is based on 11 dynamic target effect sizes and 8 static target effect sizes; two static target series (105 and 201) had extremely small sample sizes (2 and 3 sessions, respectively). A similar procedure is used in the analyses of sender/receiver pairing and experienced versus novice subjects.

[7] Three series involving laboratory senders were eliminated from this analysis because of extremely small sample sizes. These include Series 2 ($n = 2$), Series 105 ($n = 2$), and Series 201 ($n = 1$). Thus, the point biserial correlation is based on 11 series with friends and 8 series with laboratory senders.

Ganzfeld Experience

Two hundred and eighteen participants had their first experience as ganzfeld receivers in the autoganzfeld series. (This includes the 5 Novice Series 101–105 and 12 novices in Series 1.) For all but 24 (11%), their initial autoganzfeld session provided their first experience as participant in any parapsychological research. Of the 218 novices, 71 (32.5%, h = .17) correctly identified their target (exact binomial p = .0073, z = 2.44).

Participants with some ganzfeld experience contributed 137 trials and 51 hits (37%, h = .26, p = .001, z = 3.09). When series effect sizes are used as the unit of analysis and prior ganzfeld experience is used as the predictor variable, r_p is .078, t (10) = 0.25, p = .41. The 95% CI for novices is a hit rate from 25.5% to 49.5%. The CI for experienced participants is from 29% to 50%.

Participation by PRL Laboratory Staff

For completeness, we report the contribution of laboratory staff as subjects in this database. PRL staff members contributed 12 sessions as R. These sessions yield 3 hits (exact binomial p = .50; h = .00).

White Noise and Ganzfeld Illumination Levels

The mean white noise level (in arbitrary units of 0–7.5) is 2.97 (SD = 1.77). As measured from the headphones, the mean noise level is approximately 68 dB. The mean light intensity (arbitrary units of 0–100) is 73.8 (SD = 26.1). Preferred noise and light intensity levels are highly correlated: r = .569, t (353) = 12.99.

Neither noise nor light intensity is significantly related to ganzfeld performance. The point-biserial correlation between hits and noise level is −.026, t (353) = −0.48, p = .631, two tailed. For light intensity, r_p is −.040, t (353) = −0.76, p = .449, two tailed.

Randomness Tests

The adequacy of randomization was a major source of disagreement in two meta-analytic reviews of earlier psi ganzfeld research (Honorton, 1985; Hyman, 1985). In this section we document the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

adequacy of our randomization procedure according to guidelines agreed on by Hyman and Honorton (1986).

Global Tests of Random Number Generator

Full-range frequency analysis. As described earlier, autoganzfeld targets are selected through a program call to the RNG for values within the target range (1–160). The number of experimental sessions (N = 355) is too small to assess the RNG output distribution for the full range, so we performed a large-scale control series to test the distribution of values. Twelve control samples were collected. These included five samples with 156,000 trials, six samples with 1,560 trials, and one sample of 1,560,000 trials. The 12 resulting chi-square values were compared to a chi-square distribution with 155 df, using the Kolmogorov-Smirnov (KS) one-sample test. The KS test yields a two-tailed p = .577, indicating that the RNG used in these experiments provides a uniform distribution of values throughout the full target range.[8]

Test of frequency distribution for Set 20. We used a single target set (Set 20) in Series 302. We repeated the frequency analysis in a 40,000-trial control sample, restricting target selection to the four target values within Set 20 (Targets 77–80). A chi-square test of the distribution of targets within Set 20 shows that the RNG produces a uniform distribution of the target values within the set: χ^2 = 3.19, 3 df, p = .363.

Tests of the Experimental RNG Usage

Each autoganzfeld session required two RNG calls. An RNG call at the beginning of the session determined the target; another, made before the judging procedure, determined the order in which the target and decoys were presented for judging.

Distribution of targets in the experiment. A chi-square test of the distribution of values within the target sets shows that the targets were selected uniformly from among the four possibilities within each set; χ^2 with 3 df is 0.86, p = .835.

Distribution of judging order. A chi-square test of the judging order indicates that the targets were uniformly distributed among the four possible judging sequences: the χ^2 with 3 df is 1.85, p = .604.

[8] One of the preview pack elements for Set 6, containing Targets 21–24, was damaged. This required filtering the RNG calls in the experiment and control tests to bypass the damaged portion of the videotape, leaving the targets in Pool 6 unused. Thus, for the full-range analyses reported here, there are 155 df rather than 159.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Summary

The randomness tests demonstrate that the RNG used for target selection in these experiments provides an adequate source of random numbers and was functioning properly during the experiments.

EXAMPLES OF TARGET-MENTATION CORRESPONDENCES

In this section, we present some examples of correspondences between targets and ganzfeld mentation. Although conclusions cannot be drawn from qualitative data, this material should not be ignored. It constitutes the raw data on which the objective statistical evidence is based, and may provide important insights concerning the underlying process. These examples are excerpts from sessions of subjects' ganzfeld mentation reports, identified by them during the blind judging procedure as providing their basis for rating the target.

Target 90, Static: Dali's "Christ Crucified."
Series 1. Participant ID: 77. Rank = 1. z score = 1.67.

> "...I think of guides, like spirit guides, leading me and I come into like a court with a king. It's quiet.... It's like heaven. The king is something like Jesus. Woman. Now I'm just sort of summersaulting through heaven.... Brooding.... Aztecs, the Sun God.... High priest.... Fear.... Graves. Woman. Prayer.... Funeral.... Dark. Death.... Souls.... Ten Commandments. Moses...."

Target 77, Dynamic: Tidal wave engulfing ancient city. From "The Clash of the Titans," a film based on Greek mythology. A huge tidal wave crashes into the shore. The scene shifts to a center courtyard of an ancient Greek city; there is a statue in the center, and buildings with Greek columns around the periphery. People are running to escape consumption by the tidal wave. Water rushes through the buildings, destroying the columns and the statue; people scurry through a stone tunnel, just ahead of the engulfing water; debris floats through the water.
Series: 1. Participant ID: 87. Rank = 1. z score = 1.42.

> "...The city of Bath comes to mind. The Romans. The reconstruction of the baths through archaeology. The Parthenon. Also getting sort of buildings like Stonehenge but sort of a cross between Stonehenge and the Parthenon. The Byzantine Empire. The Gates of Thunder. The Holy See. Tables floating about.... The number 7 very clearly. That just popped out of nowhere. It reminds me a bit of one of the first *Clash* albums, however. The Clash, "*Two Sevens*" I think it was called, I'm not sure...." [The target was number 77.]

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Series 302. Participant ID: 267. Rank = 1. z score = 2.00.

> "...A big storm over New York City. I'm assuming it's New York City. No, it's San Francisco....A big storm and danger. It looks so beautiful but I'm getting the sense of danger from it....It's a storm. An earthquake...."

Target 63, Dynamic: Horses. From the film, "The Lathe of Heaven." An overhead view of five horses galloping in a snow storm. The camera zooms in on the horses as they gallop through the snow. The scene shifts to a close-up of a single horse trotting in a grassy meadow, first at normal speed, then in slow-motion. The scene shifts again; the same horse trotting slowly through empty city streets.
Series: 101. Participant ID: 92. Rank = 1. z score = 1.25.

> "...I keep going to the mountains....It's snowing....Moving again, this time to the left, spinning to the left....Spinning. Like on a carousel, horses. Horses on a carousel, a circus...."

Target 46, Dynamic: Collapsing Bridge. Newsreel footage of the collapse of a bridge the 1940s. The bridge is swaying back and forth and up and down. Light posts are swaying. The bridge collapses from the center into the water.
Series: 101. Participant ID: 135. Rank = 1. z score = 1.94.

> "...Something, some vertical object bending or swaying, almost something swaying in the wind....Some thin, vertical object, bending to the left....Some kind of ladder-like structure but it seems to be almost blowing in the wind. Almost like a ladder-like bridge over some kind of chasm that's waving in the wind. This is not vertical this is horizontal....A bridge, a drawbridge over something. It's like one of those old English type bridges that opens up from either side. The middle part comes up. I see it opening. It's opening. There was a flash of an old English stone bridge but then back to this one that's opening. The bridge is lifting, both sides now. Now both sides are straight up. Now it's closing again. It's closing, it's coming down, it's closed. Arc, images of arcs, arcs, bridges. Passageways, many arcs. Bridges with many arcs....

Target 137, Static: "Working on a Watermelon Farm." This painting shows a black man with his back to the picture; his suspenders form a V-shape

around his shoulders. A dog is in front of the man; there are watermelons between the dog and the man. The man faces a dirt path with watermelon patches on either side. On the left side, another man pushes a wheelbarrow filled with huge watermelons.
Series: 101. Participant ID: 105. Rank = 2. z score = 0.98.

> "...a small lamb, very soft, outside. Small, playful....I see a 'V' shape....An apple....I see a kitchen towel with a picture on it. Apple seeds or a fruit cut in half showing the seeds. A tomato or an apple. The fruit was red on the outside....I thought of watermelon as in a watermelon basket. Thinking of kids playing on a beach. Little kids playing with balls that are bigger than they are and buckets that are three-quarters their size....I had a thought of going through a tunnel, not the kind of tunnel you see on Earth but the type of tunnel described when someone dies."

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Target 64, Dynamic: 1920s Car Sinking. From the film "Ghost Story." The scene depicts the murder of a young blonde woman by three young men in the 1920s. The men are all wearing suits; one of the men is wearing a fedora hat that is turned up in the back. The men push an old car into a lake. The camera shifts between close-ups of their facial expressions, and the car, as it slowly sinks into the water. The woman's face and hand appear in the car's large rectangular rear window; she silently screams out for help. The car disappears beneath the water as the sequence ends.
Series: 102. Participant ID: 154. Rank = 1. z score = 1.45.

> "...Girl with a haircut....Blond hair....A car....The back of someone's head....Someone running to the right....Someone on the right in a brown suit...and a fedora hat turned up very much in the back....Fedora, trench coat, dark tie....A tire of a car. The car's going to the left. An old movie....I'm picturing an Edward G. Robinson movie....Big roundish car like 1940's. Those scenes from the back window. Bumping once in a while up and down looking through the back window you could see that it was probably a big screen in back of the car and the car's standing still actually....I think it's a movie I saw. They're being shot at and shooting at the window and then the girl gets shot....Girl with the blonde haircut....Someone walking in a suit, brown suit....It's the 1940's again, 30's maybe. Except it looks like it's in color. Something red, blood...blood on someone's lap....A dead person all of a sudden....A big mouth opened. Yelling, but no sound....Two people running near a train....Dressed in 1920 type suits with balloony pants, like knickers....A big, old-fashioned white car with a flat top. 1920's, 30's...."

Target 107, Static: Stained-Glass Madonna with Child. This is a stained-glass window depicting the Virgin Mary and Christ child.

Series: 102. Participant ID: 183. Rank = 2. z score = 0.61.

> "Some kind of a house, structure....Some kind of wall or building. Something with the sky in the background. Thinking of a bell. A bell structure. Something with a hole with the light coming through the hole....Like a stained glass window like you see in churches."

Target 19, Static: Flying Eagle. An eagle with outstretched wings is about to land on a perch; its claws are extended. The eagle's head is white and its wings and body are black.
Series: 104. Participant ID: 316. Rank = 1. z score = 2.00.

> "...A black bird. I see a dark shape of a black bird with a very pointed beak with his wings down....Almost needle-like beak....Something that would fly or is flying...like a big parrot with long feathers on a perch. Lots of feathers, tail feathers, long, long, long....Flying, a big huge, huge eagle. The wings of an eagle spread out....The head of an eagle. White head and dark feathers....The bottom of a bird...."

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Target 144, Dynamic: Hell. From the film "Altered States." This sequence depicts a psychedelic experience. Everything is tinted red. The rapidly shifting scenes include: A man screaming; many people in the midst of fire and smoke; a man screaming in an isolation tank; people in agony; a large sun with a corona around it; a mass crucifixion; people jumping off a precipice, in the midst of fire, smoke, and molten lava; spiraling crucifixes. There is a close-up of a lizard's head, slowly opening its mouth, at the end of the sequence.
Series: 104. Participant ID: 321. Rank = 1. z score = 1.49.

> "...I just see a big 'X'. A big 'X'....I see a tunnel in front of me. It's like a tunnel of smog or a tunnel of smoke. I'm going down it....I'm going down it at a pretty fast speed....I still see the color red, red, red, red, red, red, red, red....Ah, suddenly the sun....The kind of cartoon sun you see when you can see each pointy spike around the sphere....I stepped on a piece of glass and there's a bit of blood coming out of my foot....A lizard, with a big, big, big head...."

Target 148, Static. Three Unusual Planes. Three small aircraft flying in formation. The planes are white and have swept-back wings; their landing-gear is extended. A winding road is visible below.
Series: 104. Participant ID: 322. Rank = 2. z score = 0.39.

> "...A jet plane....A 747 on the way to Greece. Blue skies. Sounds like it's going higher....I think I'm back on the plane again. I never used to be afraid of flying until recently....They need better insulated jets, soundproof like these rooms. They could use these comfortable seats, too. And the leg room. The service isn't bad either....Still can't get the

feeling of being in an airplane out of my mind. Flying over Greenland and Iceland when I went to England.... Feels like we're going higher and higher.... Descending. It seems we're descending.... Big airplanes flying over with people like me staring down.... Flying around in a piece of tin.... Feel like I'm getting a G-force. Maybe I am taking off. Sure feels like it. Feels like we're going straight up.... I always feel like when I'm on the plane going home, I just hope that plane makes it past the Rocky Mountains...."

Target 10, Static: Santa and Coke. This is a Coca-Cola Christmas ad from the 1950s, showing Santa Claus holding a Coke bottle in his left hand; three buttons are visible on Santa's suit. Behind Santa and to his left, is a large bottle cap with the Coca-Cola logo leaning against an ornamented Christmas tree.
Series: 104. Participant ID: 332. Rank = 1. z score = 1.14.

"... There's a man with a dark beard and he's got a sharp face.... There's another man with a beard. Now there's green and white and he's in bushes and he's sort of colonial. He looks like Robin Hood and he's wearing a hat.... I can see him from behind. I can see his hat and he has a sack over his shoulder.... Window ledge is looking down and there's a billboard that says 'Coca-Cola' on it.... There's a snowman again and it's got a carrot for a nose and three black buttons coming down the front.... There's a white beard again. There's a man with a white beard.... There's an old man with a beard...."

Target 70, Dynamic: Dancing in NY City Streets. From the film "The Wiz." The span of yellow-paved bridge over a body of water and automobile traffic is visible in the opening scene; the New York City skyline is in the background. A hot-air balloon flies overhead. The scene shifts as Dorothy (Diana Ross), her dog Toto, the Lion, Tin Man, and Scarecrow dance along the bridge; one of the bridge's supporting arches is behind them. The Chrysler Building is in the background. At the end of the sequence, the characters dance in front of a painted backdrop of an old-fashioned building.
Series: 105. Participant ID: 336. Rank = 1. z score = 1.40.

"Big colorful hot air balloons.... White brick wall.... Ocean.... People walking before my eyes. Several people.... A dog. Hot air balloon.... a nightclub singer.... Back of a woman's head, short curly hair.... Water.... Balloon, big balloon.... Yellow.... Very tall building. Looking down at a city. Leaving a city, going up.... Faces. An arc.... Water.... A woman's face.... Cars, freeway.... A rock-n-roll star chanting.... Architecture. A jester's hat.... geometrical figures, designs. ... Yellow chocolate bar. Water. Going down into water, deep down.... Man with long golden hair and sun glasses.... The Bay, San Francisco

Bay. A lion.... Highways.... Lion, see a lion.... Tornado.... Balloon.... Face mask.... City.... Leaning Tower of Pisa.... Long hallway, doorway.... Long road. Long, long desert road...."

Target 22, Dynamic: Spiders. From the documentary "Life on Earth." A spider is weaving its web. The spider's long legs spring up and down repeatedly, weaving strands of the web. The body of the spider is constantly in motion, and bounces up and down. A close-up shows one of the veins of the web being stretched out by the spider. Various views of the web.
Series: 301. Participant ID: 146. Rank = 2. z score = 0.65.

"... Now visual patterns more like a spider web and the color. And then like the form of the veins of a windmill.... Something like a spider web again. A spider web. A pattern that instead of a spider web it looks like basket weaving.... An image of the way some children were able to do something like flying when I was a child though I never had one. It was a—forgotten what it was called—a pogo stick or a jump stick, something in which you jumped up and down and you could hop quite a distance by doing so.... I have kinesthetic images all over as in vigorous motion expressed in flying or jumping on this sort of spring stick that I mentioned.... Vigorous motion. It's as though I were trying to combine relaxation with participating in an image of something very vigorous.... I really feel carried away by these images of vigorous activity without being able to localize this activity as to what it is...."

Target 108, Static: Two fire eaters. A young fire eater, in the foreground, facing to the right of the picture, blows a huge flame out of his mouth. In the background there is another fire eater. A group of people are watching on the left side of the picture.
Series: 301. Participant ID: 146. Rank = 1. z score = 1.71.

"... I keep having images of flames now and then.... The sound reminds me of flames too.... I find flames again.... In these new images the fire takes on a very menacing meaning.... Rather mountainous sticking up of bare rocks just as though they had come from a recently formed volcano. Volcanos of course get back to the fire, extreme heat. I had an image of a volcano with molten lava inside the crater. Molten lava running down the side of the volcano.... Cold. Written out there behind the visual field and thinking how it contrasts with my images of flames. Although my images of flames didn't actually include much real feeling of heat. I didn't have any imagery of heat in connection with the flames. Just abstract thought of flames.... Now I think of the water as a way of putting out flames. Suddenly, I was biting my lip. Biting my lip as though lips had something to do with the imagery and I see lips out in front of me.... And the lips I see are bright red, reminding me of the flame imagery earlier. And then a bright heart such as Valentine's

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4
Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

candy in the shape of a heart. The cinnamon flavored candies that I remember as a child having at Valentine's. Red color.... This red as in the cinnamon candy is a deep very intense red. And similarly for the flames. And now I see the word 'red'...."

Target 94, Dynamic: Hang Gliders. The sequence shows a skier on a V-shaped hang glider. The skier soars high up above snow covered mountains and a pine forest. At the end, the skier lands on a mountain slope and skis away. The sequence is accompanied by Pachelbel's Canon.
Series: 301. Participant ID: 188. Rank = 1. z score = 1.26.

"...Some kind of 'V' shape, like an open book.... I get some mountain.... Some kind of bird with a long wing.... The shape of an upside down 'V'.... Ski, something about skiing came to me.... Some kind of a body like an oval shape of a body with wings on top of it in a 'V' shape. Another 'V' like a wing shape.... Something with wings.... Again the shape of an umbrella came into my mind. A butterfly shape...."

Target 80, Dynamic: Bugs Bunny in Space. In this cartoon, there is a close-up of the lower part of a cigar-shaped rocketship and the supports holding it up. The rocket assembly slides over to the launching pad, directly above Bugs Bunny's underground patch. The scene shifts to the underground patch, as Bugs Bunny climbs up the ladder leading out of his patch. Unknowingly, he climbs up through the interior of the rocketship. The rocket's supports pull away and then it takes off into space. The rocket's nose cone spins as Bugs Bunny appears through the top and he sees the Earth recede rapidly in the distance. As the sequence ends, Bugs Bunny is hit in the belly by a comet.
Series: 302. Participant ID: 292. Rank = 1. z score = 1.48.

"...Space craft.... The solar system. The underside of a helicopter or a submarine or some kind of fish that you're seeing from underneath.... Sort of being underneath it. Sort of being underneath it.... A very strange image like a cartoon character, animated character. With his mouth open kind of.... Like a hypodermic needle or a candle or this shaft like thing with the a pointed top again.... missiles flying.... An aerial perspective.... I'm just kind of editing here I think. I'm really hoping all this rocketship kind of imagery isn't because of the noise. I feel like I'm in a rocketship or something.... That image of the ship going into the belly of the mother ship...."

Comparison of Study Outcomes with Ganzfeld Meta-Analysis

In this section, we compare the automated ganzfeld study outcomes with the results of earlier ganzfeld studies, summarized in a

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 5
Comparison of Overall Performance in Automated Ganzfeld and Meta-Analysis Data Sets

Outcome variable	Database	N studies	Mean	SD	t	df	p
z scores	Meta-analysis	28	1.25	1.57	0.33	25	.748
	Autoganzfeld	11	1.10	1.14			
Effect sizes (h)	Meta-analysis	28	.28	.46	0.14	28	.892
	Autoganzfeld	11	.29	.29			

Note. The p values are two-tailed.

meta-analysis (Honorton, 1985). We compare the two databases on four dimensions: (1) overall success rate, (2) dynamic versus static targets, (3) sender/receiver pairing, and (4) novice versus experienced subjects.

Overall Success Rate

To assess the consistency of results, we compare the 11 autoganzfeld series to the 28 studies in a meta-analysis of earlier ganzfeld studies (Honorton, 1985, Table A1, p. 84), using direct hits as the dependent variable. The outcomes of the two data sets are consistent. Both display a predominance of positive outcomes: 23 of the 28 studies in the meta-analysis (82%) and 10 of the 11 autoganzfeld series (91%) yield positive z scores. The mean autoganzfeld z scores and effect sizes are very similar to those in the meta-analysis. (See Table 5.)

Combined Estimates of Ganzfeld Success Rate

Because the z scores and effect sizes for the automated ganzfeld are consistent with the original set of 28 studies in the meta-analysis, a better estimate of their true population values may be obtained by combining them. Positive outcomes were obtained in 33 of the 39 studies (85%); the 95% CI is from 69% to 99%. Table 6 shows a stem-and-leaf frequency plot of the z scores (Tukey, 1977). Unlike other methods of displaying frequency distributions, the stem-and-leaf plot retains the numerical data precisely. (Turned on its side, the stem-and-leaf plot becomes a conventional histogram.) Each number includes a stem and one or more leaves. For example, the stem 1 is followed by leaves of 6,6,6,7,7,7, representing z scores of 1.6,1.6,1.6,1.7,1.7,1.7. In the display, the letter "H" identifies the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 6
DISTRIBUTION OF Z SCORES

Stem		Leaf			
			Minimum z	=	−1.97
−1.		97	Lower hinge	=	0.25
−0.		85	Median z	=	0.92
−0.		33	Mean z	=	1.28
0.	H	222224	Upper hinge	=	2.08
0.	M	6667777999	Maximum z	=	4.02
1.		666777	*SD*	=	1.44
2.	H	011	Skewness (g_1)	=	0.05
2.		8	Kurtosis (g_2)	=	−0.37
3.		01124	Combined (Stouffer) z	=	7.53
3.		9			
4.		0			

upper and lower hinges of the distribution, and "M" identifies its median. The z's range from −1.97 to 4.02 (mean z = 1.21, *SD* = 1.45), and the 95% CI is a z from .76 to 1.66.

The combined z for the 39 studies is 7.53 ($p = 9 \times 10^{-14}$). Rosenthal's (1984) file-drawer statistic indicates that 778 additional studies with z scores averaging zero would be required to reduce the significance of the combined ganzfeld database to nonsignificance; that is a ratio of 19 unknown studies for every known study.

A stem-and-leaf display of the effect sizes is shown in Table 7. The effect sizes range from −.93 to 1.44 (mean *h* = .28, *SD* = .41). The two most extreme values on both sides of the distribution are outliers. The 95% CI is an *h* between .15 and .41; the equivalent hit rate is from 31.5% to 44.5%.

Dynamic Versus Static Targets

The use of video sequences as targets is a novel feature of the autoganzfeld database. However, a comparable difference in target type exists in the earlier ganzfeld studies. Of the 28 direct hits studies in the meta-analysis, 9 studies (by three independent investigators) used *View Master* stereoscopic slide reels as targets (Honorton, 1985, Studies 7–8, 16–19, 21, 38–39). Static targets (single pictures or slides) were used in the remaining 19 studies by seven independent investigators (Studies 1, 2, 4, 10–13, 23–31, 33–34, 41–42). Like the autoganzfeld video sequences, *View Master* targets present a variety of images reinforcing a central target theme.

TABLE 7
DISTRIBUTION OF EFFECT SIZES (COHEN'S *h*)

Stem		Leaf			
−.9		3			
−.4		0			
OUTSIDE VALUES					
			Minimum *h*	=	−0.93
−.3		1	Lower hinge	=	0.10
−.1		0	Median *h*	=	0.25
−.0		51	Mean *h*	=	0.28
.0		7779	Upper hinge	=	0.41
.1	H	002888	Maximum *h*	=	1.44
.2	M	1334	*SD*	=	0.41
.3		11144777	Skewness (g_1)	=	0.28
.4	H	01113	Kurtosis (g_2)	=	2.49
.5		7			
.7		3			
.8		17			
OUTSIDE VALUES					
1.3		3			
1.4		4			

To compare the relative impact of dynamic and static targets in the autoganzfeld and meta-analysis, we obtained point-biserial correlations for each data set using target type (static or dynamic) as the predictor variable and the series effect size, Cohen's *h*, as the outcome variable. We test the difference between the two correlations using Cohen's *q* (Cohen, 1977). Dynamic targets yield significantly larger effect sizes in both data sets. For the meta-analysis, r_p is .409, *t* (26) = 2.28, *p* = .015; and for the autoganzfeld, as reported above, r_p is .663. The two correlations are not significantly different (*q* = .36; z = 1.14). Therefore, we combine the two data sets to obtain a better estimate of the relationship between effect size and target type: r_p = .439, *t* (45) = 3.28, *p* = .002. The 95% CIs are 24% to 36% for static targets and 38% to 55% for dynamic targets. Thus, the cumulative evidence strongly indicates that dynamic targets are more accurately retrieved than static targets.

Sender/Receiver Pairing

A similar analysis compares the effects of sender/receiver pairing in the two databases. Studies in the meta-analysis did not routinely

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

provide detailed breakdowns regarding sender/receiver pairing. Sender/receiver pairing in the meta-analysis can only be coded according to whether subjects could bring friends to serve as their sender or were restricted to laboratory senders. In 17 studies, by six independent investigators, subjects were free to bring friends (Honorton, 1985, Studies 1–2, 4, 7–8, 16, 23–28, 30, 33–34, 38–39). Laboratory-assigned senders were used exclusively in the remaining 8 studies, by four independent investigators (Studies 10–12, 18–19, 21, 29, 41). (Three studies using clairvoyance procedures and no senders are excluded from this analysis.) For the autoganzfeld studies, we calculated separate effect sizes for each series by sender type (combining lab friend and friend for comparability with the meta-analysis). In the meta-analysis, r_p (23) is .403; larger effect sizes occurred in studies where friends could serve as sender (t = 2.11, p = .023). For the autoganzfeld, as reported above, r_p is .363, in the same direction. The two correlations are very similar (q = .05; z = 0.14) and are combined to give a better estimate of the relationship between sender/receiver pairing and ganzfeld study outcome: r_p = .38, t (42) = 2.66, p = .0055. The 95% CIs are 20% to 34% for unacquainted sender/receiver pairs and 34.1% to 49.2% for friends. Thus, the sender/receiver relationship does have a significant impact on performance.

Effect of Prior Ganzfeld Experience

The meta-analysis includes 14 studies, by nine independent investigators, in which novices are used exclusively (Honorton, 1985, Studies 2, 4, 8, 10–12, 16–18, 23–24, 31, 41–42). Experienced or mixed samples of novice and experienced subjects are used in the remaining 14 studies, by four different investigators (Studies 1, 7, 19, 21, 25–30, 33–34, 38–39). Studies using experienced subjects were more successful than those limited to novices; the point-biserial correlation between level of experience and effect size is .229, t (26) = 1.20, p = .12. For the autoganzfeld studies, as reported above, r_p is .078. The two correlations do not differ significantly (q = .155; z = 0.40), and the combined r_p is .194, t (38) = 1.22, p = .105. The respective 95% CIs are 24.5% to 44.5% for novices and 35.5% to 48% for experienced subjects.

The 95% CIs for these comparative analyses are shown graphically in Figure 2. The bottom two rows are CIs for the overall hit rates in the meta-analysis and autoganzfeld, respectively. The next

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

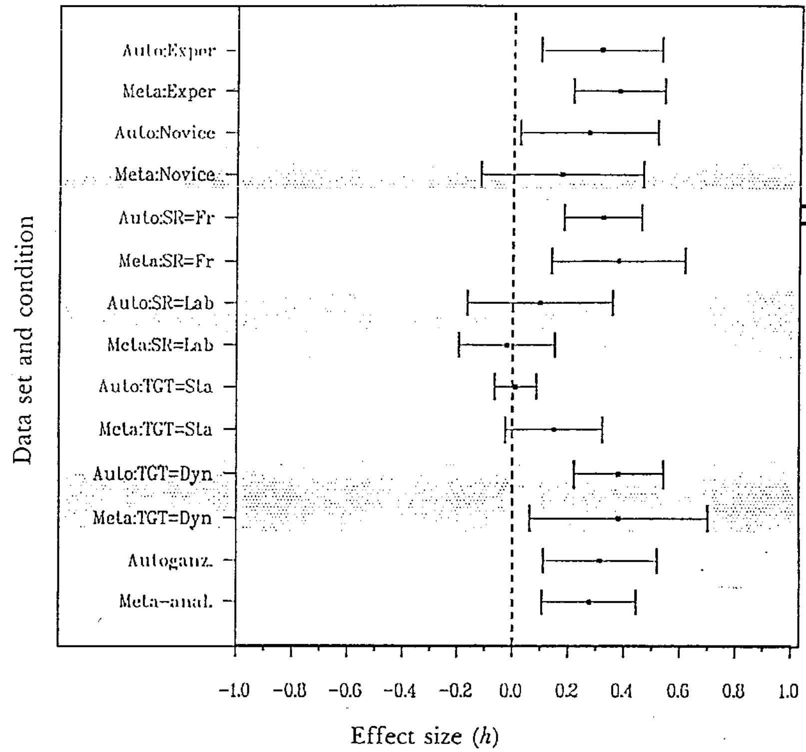


Figure 2. Comparison of autoganzfeld and meta-analysis 95% confidence limits. Abbreviations are defined as follows: Meta = meta-analysis studies, Auto = automated ganzfeld studies, Dyn = dynamic targets, Sta = static targets, Lab = laboratory senders, Fr = sender is friend or acquaintance of receiver, Novice = no prior ganzfeld experience, Exper = prior ganzfeld experience.

two rows give the CIs for dynamic targets in the two data sets, and so on.

Discussion

We now consider various rival hypotheses that might account for the experimental outcomes, and the degree to which the automated ganzfeld experiments, viewed in conjunction with the earlier psi

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

ganzfeld studies, constitute evidence for psi communication. Finally, we consider directions for future research suggested by these findings.

Rival Hypotheses

Sensory Cues. Only Se knows the identity of the target until R finishes the automated judging procedure. If Se is not a PRL staff member, a staff member not otherwise involved in the session supervises target selection. In either case, the target selector knows only which videocassette contains the target. The target selector leaves the monitoring room with the remaining three target tapes after knocking three times on the monitoring room door, signalling E to return. Since the target selector only knows the videocassette number, variations in knocking cannot communicate any useful information to E. The cardboard cover over the VCR eliminates any visual cues to E regarding the position of the videotape or the activity of the VU meters (which are active when the target is dynamic and has a soundtrack).

Sensory transmission from Se to R during the ganzfeld session is eliminated by having R and Se in separate, sound-attenuated rooms. If either participant leaves their room before R's ratings have been registered in the computer, the session is unconditionally aborted.

The videotape target display system prevents potential handling cues during the judging procedure. Computer registration of R's target ratings and automated feedback after the session prevents the possibility of cheating by Se during feedback, raised by Hyman (1985).

After about 80% of the sessions were completed, it was becoming clear that our hypothesis concerning the superiority of dynamic targets over static targets was receiving substantial confirmation. Because dynamic targets contain auditory as well as visual information, we conducted a supplementary test to assess the possibility of auditory leakage from the VCR soundtrack to R. With the VCR audio set to normal amplification, no auditory signal could be detected through R's headphones, with or without white noise. When an external amplifier was added between the VCR and R's headphones and with the white noise turned completely off, the soundtrack could sometimes be faintly detected. It is unlikely that subjects could have detected any target audio signal with the normal VCR amplification and white noise; as we have reported, there is no correlation between ganzfeld success rate and white noise level in these experiments. Nevertheless, to totally exclude any possibility of subliminal cueing, we modified the equipment. Additional testing confirmed that this modification effectively eliminated all leakage. This was formally confirmed by an audio spectrum analysis, covering the frequency domain between 475 Hz and 15.2 kHz. The critical question, of course, is whether performance on dynamic targets diminished after this modification. The answer is no; in fact, performance improved. Before the modification, the direct hit rate on dynamic targets was 38% (150 trials, 57 hits, $h = .28$, exact binomial $p = .00029$, $z = 3.44$); the 95% CI was from 31% to 45%. Following the modification, the direct hit rate was 50% (40 trials, 20 hits, $h = .52$, exact binomial $p = .00057$, $z = 3.25$) with a 95% CI from 37% to 63%. The direct hit rate for all targets—static and dynamic—after the modification was 44% (64 trials, 28 hits, $h = .39$, exact binomial $p = .00082$, $z = 3.15$).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Randomization. As Hyman and Honorton (1986, p. 357) have pointed out, "Because ganzfeld experiments involve only one target selection per session . . . , the ganzfeld investigator can restrict his or her attention to a frequency analysis allowing assessment of the degree to which targets occur with equal probability." We have documented both the general adequacy of the RNG used for target selection and its proper functioning during the experiment.

Data selection. Except for two pilot studies, the number of participants and trials were specified in advance for each series. The pilot or formal status of each series was similarly specified in advance and recorded on disk before beginning the series. We have reported all trials, including pilot and ongoing series, using the automated ganzfeld system. Thus, there is no "file-drawer" problem in this database.

Psi ganzfeld success rate is similar for pilot and formal sessions. The proportion of hits for the 66 pilot sessions is .32 ($h = .16$, $p = .129$, $z = 1.13$). For the 289 formal sessions, the proportion correct is .35 ($h = .22$, $p = .0001$, $z = 3.71$). The difference is not significant: $\chi^2 = 0.11$, 1 *df*, $p = .734$.

If we assume that the remaining trials in the three unfinished series would yield only chance results, these series would still be statistically significant (exact binomial $p = .009$, $z = 2.36$). This would reduce the overall z for all 11 series from 3.89 to 3.61. Thus, inclusion of the three incomplete studies does not pose an optional stopping problem.

Multiple analysis. Informal examination of recent issues of several American Psychological Association journals suggests that correction for multiple comparisons is not a common practice in more conventional areas of psychological inquiry. Nevertheless, half of Hyman's (1985) 50-page critique of earlier psi ganzfeld research focused on issues related to multiple testing. In the present case, advance specification of the primary hypothesis and method of analysis prevents problems involving multiple analysis or multiple indices in our test of the overall psi ganzfeld effect. Our direct hits analysis is actually *less* significant than either the sum of ranks method ($z = 4.04$, $p = 2.7 \times 10^{-5}$) or Stanford's z scores ($t = 4.53$, 354 *df*, $p = 4.1 \times 10^{-6}$).

In addition to the primary hypothesis, however, we also tested two secondary hypotheses concerning the impact of target type and sender/receiver pairing on psi performance, and we have presented several purely exploratory analyses as well. Our Results section includes 15 significance tests involving psi performance as the dependent variable, and the p values cited are not adjusted for multiple comparisons. Of the 15 significance tests, 9 are associated with $p < .05$. The Bonferroni multiple comparisons procedure provides a conservative method of adjusting the alpha level when several simultaneous tests of significance are performed (Holland & Copenhaver, 1988; Hyman & Honorton, 1986; Rosenthal & Rubin, 1984). When the Bonferroni adjustment is applied, six of the nine individually significant outcomes remain significant; these are: the overall hit rate, the subject-based analysis using Stanford z scores, the difference between dynamic and static targets, the dynamic target hit rate, and the hit rate for experienced subjects.

Although the relationship between psi performance and sender type is not independently significant in the autoganzfeld, the correlation coefficient of .363 is close to that observed in the meta-analysis ($r = .403$), and the combined result is significant. The cumulative evidence, therefore, does support the conclusion that the sender/receiver relationship is a significant moderator of ganzfeld psi performance.

Security. Given the large number of subjects and the significance of the outcome using subjects as the unit of analysis, subject deception is not a plausible explanation. The automated ganzfeld protocol has been examined by several dozen parapsychologists and behavioral researchers from other fields, including well-known critics of parapsychology. Many have participated as subjects, senders, or observers. All have expressed satisfaction with our handling of security issues and controls.

In addition, two experts on the simulation of psi ability have examined the autoganzfeld system and protocol. Ford Kross has been

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

a professional mentalist for over 20 years. He is the author of many articles in mentalist periodicals and has served as Secretary/Treasurer of the Psychic Entertainers Association. Mr. Kross has provided us with the following statement: "In my professional capacity as a mentalist, I have reviewed Psychophysical Research Laboratories' automated ganzfeld system and found it to provide excellent security against deception by subjects" (personal communication, May, 1989). We have received similar comments from Daryl Bem, Professor of Psychology at Cornell University. Professor Bem is well known for his research in social and personality psychology. He is also a member of the Psychic Entertainers Association and has performed for many years as a mentalist. He visited PRL for several days and was a subject in Series 101.

The issue of investigator integrity can only be conclusively addressed through independent replications. It is, however, worth drawing attention to the 13 sessions in which a visiting scientist, Marilyn J. Schlitz, served as either experimenter ($N = 7$, 29% hits, $h = .08$) or sender ($N = 6$, 67% hits, $h = .36$). Altogether, these sessions yielded 6 direct hits ($N = 13$, 46.2% hits, $h = .45$). This effect size is more than twice as large as that for the database as a whole.

Status of the Evidence for Psi Communication in the Ganzfeld

The automated ganzfeld studies satisfy the methodological guidelines recommended by Hyman and Honorton (1986). The results are statistically significant. The effect size is homogeneous across 11 experimental series and eight different experimenters. Moreover, the autoganzfeld results are consistent with the outcomes of the earlier, nonautomated ganzfeld studies; the combined z of 7.53 would be expected to arise by chance less than one time in 9 trillion.

We have shown that, contrary to the assertions of certain critics (Druckman & Swets, 1988, p. 175), the ganzfeld psi effect exhibits "consistent and lawful patterns of covariation found in other areas of inquiry." The automated ganzfeld studies display the same patterns of relationships between psi performance and target type, sender/receiver acquaintance, and prior testing experience found in earlier ganzfeld studies, and the magnitude of these relationships is consistent across the two data sets. The impact of target type and sender/receiver acquaintance is also consistent with patterns in spontaneous case studies, linking ostensible psi experiences to emotionally significant events and persons. These findings cannot be ex-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

plained by conventional theories of coincidence (Diaconis & Mosteller, 1989).

Hyman and Honorton (1986) have stated,

> ...the best way to resolve the [ganzfeld] controversy...is to await the outcome of future ganzfeld experiments. These experiments, ideally, will be carried out in such a way as to circumvent the file-drawer problem, problems of multiple analysis, and the various defects in randomization, statistical application, and documentation pointed out by Hyman. If a variety of parapsychologists and other investigators continue to obtain significant results under these conditions, then the existence of a genuine communications anomaly will have been demonstrated. (pp. 353–354)

We have presented a series of experiments that satisfy these guidelines. Although no single investigator or laboratory can satisfy the requirement of independent replication, the automated ganzfeld studies are quite consistent with the earlier studies. On the basis of the cumulative evidence, we conclude that the ganzfeld effect represents a genuine communications anomaly. This conclusion will either be strengthened or weakened by additional independent replications, but there is no longer any justification for the claim made by some critics that the existing evidence does not warrant serious attention by the scientific community.

Recommendations for Future Research

Recent psi ganzfeld research has necessarily focused on methodological issues arising from the ganzfeld controversy. It is essential that future studies comply with the methodological standards agreed on by researchers and critics. Yet it is equally imperative that serious attention be given to conditions associated with successful outcomes.

Small to medium effect sizes characterize many research findings in the biomedical and social sciences (e.g., Cohen, 1977; Rosenthal, 1984). Rosenthal (1986) and Utts (1986) make a strong case for more careful consideration of the magnitude of effect in the design and analysis of future ganzfeld studies. The automated ganzfeld studies show a success rate slightly in excess of 34%. Utts's (1986) power analysis shows that for an effect of this size, the investigator has only about one chance in three of obtaining a statistically significant result in a 50-trial experiment. Even with 100 trials—an unusually large sample size in ganzfeld research—the probability of a significant outcome is only about .5.

We urge ganzfeld investigators to use dynamic targets and to design their studies to allow subjects to have the option to have friends or acquaintances as their senders. The similarity of the autoganzfeld and meta-analysis data sets strongly indicates that these factors are important moderators of psi ganzfeld performance. If our estimate of the impact of dynamic and static targets is accurate, a 50-session series using dynamic targets has approximately an 84% chance of yielding a significant outcome. A comparable series with static targets has only about one chance in five of achieving significance.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

References

Alcock, J. E. (1986). Comments on the Hyman-Honorton ganzfeld controversy. *Journal of Parapsychology*, **50**, 345–348.

Akers, C. (1984). Methodological criticisms of parapsychology. In S. Krippner (Ed.), *Advances in parapsychological research, Vol. 4* (pp. 112–164). Jefferson, NC: McFarland.

Berger, R. E., & Honorton, C. (1986). An automated psi ganzfeld testing system. In D. H. Weiner & D. I. Radin (Eds.), *Research in parapsychology 1985* (pp. 85–88). Metuchen, NJ: Scarecrow Press.

Blackmore, S. (1980). The extent of selective reporting of ESP ganzfeld studies. *European Journal of Parapsychology*, **3**, 213–219.

Blackmore, S. (1987). A report of a visit to Carl Sargent's laboratory. *Journal of the Society for Psychical Research*, **54**, 186–198.

Braud, W. G. (1978). Psi conducive conditions: Explorations and interpretations. In B. Shapin & L. Coly (Eds.), *Psi and states of awareness* (pp. 1–34). New York: Parapsychology Foundation, Inc.

Braud, W. G., Wood, R., & Braud, L. W. (1975). Free-response GESP performance during an experimental hypnagogic state induced by visual and acoustic ganzfeld techniques: A replication and extension. *Journal of the American Society for Psychical Research*, **69**, 105–113.

Briggs, K. C., & Myers, I. B. (1957). *Myers-Briggs Type Indicator Form F*. Palo Alto, CA: Consulting Psychologists Press, Inc.

Brownlee, K. A. (1965). *Statistical theory and methodology in science and engineering*. New York: John Wiley & Sons, Inc.

Child, I. L. (1986). Comments on the ganzfeld controversy. *Journal of Parapsychology*, **50**, 337–344.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

Diaconis, P., & Mosteller, F. (1989). Methods for studying coincidences. *Journal of the American Statistical Association*, **84**, 853–861.

Druckman, D., & Swets, J. (1988). *Enhancing human performance: Issues, theories, and techniques*. Washington, DC: National Academy Press.

Harley, T., & Matthews, G. (1987). Cheating, psi, and the appliance of science: A reply to Blackmore. *Journal of the Society for Psychical Research*, **54**, 199–207.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Volume 19, Number 12, December 1989

Plenum Press • New York–London

This issue completes Volume 19

FNDPA4 19(12) 1441–1538 (1989)

ISSN 0015-9018

FOUNDATIONS OF PHYSICS

An International Journal Devoted to the Conceptual Bases and Fundamental Theories of Modern Physics, Biophysics, and Cosmology

Editor:
Alwyn van der Merwe

Editorial Board

Asim O. Barut
Peter G. Bergmann
Nikolai N. Bogolubov
David Bohm
Robert S. Cohen
Olivier Costa de Beauregard
Robert H. Dicke
Bai-lin Hao
Max Jammer
Brian D. Josephson
R. Bruce Lindsay
Per-Olov Löwdin
Henry Margenau
Jagdish Mehra
André Mercier
Louis Néel
Kazuhiko Nishijima
James L. Park
Linus Pauling
Rudolph Peierls
Karl R. Popper
Ilya Prigogine
Abdus Salam
John L. Synge
Hans-J. Treder
Jean-Pierre Vigier
Mikhail Vol'kenshtein
Carl Friedrich von Weizsäcker
Eugene P. Wigner
Chen-Ning Yang

Founding Editors: Henry Margenau and Wolfgang Yourgrau†

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Foundations of Physics, Vol. 19, No. 12, 1989

Evidence for Consciousness-Related Anomalies in Random Physical Systems

Dean I. Radin[1] and Roger D. Nelson[2]

Received May 6, 1988; revised June 12, 1989

Speculations about the role of consciousness in physical systems are frequently observed in the literature concerned with the interpretation of quantum mechanics. While only three experimental investigations can be found on this topic in physics journals, more than 800 relevant experiments have been reported in the literature of parapsychology. A well-defined body of empirical evidence from this domain was reviewed using meta-analytic techniques to assess methodological quality and overall effect size. Results showed effects conforming to chance expectation in control conditions and unequivocal non-chance effects in experimental conditions. This quantitative literature review agrees with the findings of two earlier reviews, suggesting the existence of some form of consciousness-related anomaly in random physical systems.

1. INTRODUCTION

The nature of the relationship between human consciousness and the physical world has intrigued philosophers for millenia. In this century, speculations about mind–body interactions persist, often contributed by physicists in discussions of the measurement problem in quantum mechanics. Virtually all of the founders of quantum theory—Planck, de Broglie, Heisenberg, Schrödinger, Einstein—considered this subject in depth,[1] and contemporary physicists continue this tradition.[2–7]

[1] Department of Psychology, Princeton University, Princeton, New Jersey 08544. Present address: Contel Technology Center, 15000 Conference Center Drive, P.O. Box 10814, Chantilly, Virginia 22021-3808.

[2] Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, New Jersey 08544.

0015-9018/89/1200-1499$06.00/0 © 1989 Plenum Publishing Corporation

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The following expression of the problem can be found in a recent interpretation of quantum theory:

> If conscious choice can decide what particular observation I measure, and therefore into what states my consciousness splits, might not conscious choice also be able to influence the outcome of the measurement? One possible place where mind may influence matter is in quantum effects. Experiments on whether it is possible to affect the decay rates of nuclei by thinking suitable thoughts would presumably be easy to perform, and might be worth doing.[8]

Given the distinguished history of speculations about the role of consciousness in quantum mechanics, one might expect that the physics literature would contain a sizable body of empirical data on this topic. A search, however, reveals only three studies.

The first is in an article by Hall, Kim, McElroy, and Shimony, who reported an experiment "based upon taking seriously the proposal that the reduction of the wave packet is due to a mind–body interaction, in which both of the interacting systems are changed."[9] This experiment examined whether one person could detect if another person had previously observed a quantum mechanical event (gamma emission from sodium-22 atoms). The idea was based on the supposition that if person A's observation actually changes the physical state of a system, then when person B observes the same system later, B's experience may be different according to whether A has or has not looked at the system. Hall *et al.*'s results, based on a total of 554 trials, did not support the hypothesis; the observed number of "hits" obtained in their experiment was precisely the number expected by chance (277), while the variance of their measurements was significantly smaller than expected ($p < 0.05$).[9]

The second study is referred to by Hall *et al.*, who end their article by pointing out that a similar, unpublished experiment using cobalt-57 as the source was successful (40 hits out of 67 trials).[10]

The third study is a more systematic investigation reported by Jahn and Dunne,[11] who summarize results of over 25 million binary trials collected during seven years of experimentation with random-event generators. These experiments, involving long-term data collection with 33 unselected individuals, provide persuasive, replicable evidence of an anomalous correlation between conscious intention and the output of random number generators.

Thus, of three pertinent experiments referenced in mainstream physics journals, one describes results statistically too close to chance expectation and two describe positive effects.[9-11] Given the theoretical implications of such an effect, it is remarkable that no further experiments of this type can be found in the physics literature; but this is not to say that no such experiments have been performed. In fact, dozens of researchers have

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

reported conceptually identical experiments in the puzzling and uncertain domain of parapsychology. Perhaps because of the insular nature of scientific disciplines, the vast majority of these experiments are unknown to most scientists. A few critics who have considered this literature have dismissed the experiments as being flawed, nonreplicable, or open to fraud,[12–16] but their assertions are countered by at least two detailed reviews which provide strong statistical support for the existence of anomalous consciousness-related effects with random number generators.[17,18] In this paper, we describe the results of a comprehensive, quantitative meta-analysis which focused on the questions of methodological quality and replicability in these experiments.

2. THE EXPERIMENTS

The experiments involved some form of microelectronic random number generator (RNG), a human observer, and a set of instructions for the observer to attempt to "influence" the RNG to generate particular numbers, or changes in a distribution, solely by intention. RNGs are usually based upon a source of truly random events such as electronic noise, radioactive decay, or randomly seeded pseudorandom sequences.[19] Feedback about the distribution of random events is often provided in the form of a digital display, but audio feedback, computer graphics, and a variety of other mechanisms have also been used. Some of the RNGs described in the literature are technically sophisticated, the best devices employing electromagnetic shielding, environmental failsafe mechanisms triggered by deviant voltages, currents, or temperature, automatic computer-based data recording on magnetic media, redundant hard copy output, periodic randomness calibrations, and so on.[18,20]

RNGs are typically designed to produce a sequence of random bits at the press of a button. After generating a sequence of say, 100 random bits (0's or 1's), the number of 1's in the sequence may be provided as feedback. In an experimental protocol using a binary RNG, a run might consist of an observer being asked to cause the RNG to produce, in three successive button presses, a high number (sum of 1's greater than chance expectation of 50), a low number (less than 50), and a control condition with no directional intention. An experiment might consist of a group of individuals each contributing a hundred such runs, or one individual contributing several thousand runs. Results are usually analyzed by comparing high aim and low aim means against a control mean or theoretical chance expectation.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

3. META-ANALYTIC PROCEDURES

The quantitative literature review, also called meta-analysis, has become a valuable tool in the behavioral and social sciences.[21] Meta-analysis is analogous to well-established procedures used in the physical sciences to determine parameters and constants. The technique assesses replication of an effect within a body of studies by examining the distribution of effect sizes.[22-24] In the present context, the null hypothesis (no mental influence on the RNG output) specifies an expected mean effect size of zero. A homogeneous distribution of effect sizes with nonzero mean indicates replication of an effect, and the size of the deviation of the mean from its expected value estimates the magnitude of the effect.

Meta-analyses assume that effects being compared are similar across different experiments, that is, that all studies seek to estimate the same population parameters. Thus the scope of a quantitative review must be strictly delimited to ensure appropriate commonality across the different studies that are combined.[21,25] This can present a nontrivial problem in meta-analytic reviews because replication studies typically investigate a number of variables in addition to those studied in the original experiments. In the present case, because different subjects, experimental protocols, and RNGs were employed within the reviewed literature, some heterogeneity attributable to these factors was expected in the obtained distribution of effect sizes. However, the circumscription for the review required that every study in the database have the same primary goal or hypothesis, and hence estimate the same underlying effect.

Experiments selected for review examined the following hypothesis: The statistical output of an electronic RNG is correlated with observer intention in accordance with prespecified instructions, as indicated by the directional shift of distribution parameters (usually the mean) from expected values.

Because this "directional shift" is most often reported as a standard normal deviate (i.e., Z score) in the reviewed experiments, we determined effect size as a Z score normalized by the square root of the sample size (N), $e = Z/\sqrt{N}$, where N was the total number of individual random events (with probability of a hit at $p = 0.5$, $p = 0.25$, etc.). This effect size measure is equivalent to a Pearson product moment correlation.[21]

3.1. Unit of Analysis

To avoid redundant inclusion of data in a meta-analysis, "units of analysis" are often specified. We employed the following method: If an author distinguished among several experiments reported in a single

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

article with titles such as "pilot test" or "confirmatory test," or provided independent statistical summaries, each of these studies was coded and quality-assessed separately. If an experiment consisted of two or more conditions comparing different intentions or types of RNG devices, the data were split into separate units of analysis to allow the results to be coded unambiguously. In general, within a given reviewed report, the largest possible aggregation of nonoverlapping data collected under a single intentional aim was defined as the unit of analysis (hereafter called an experiment or study).

For each experiment, a Z score was assigned corresponding to whether the observed result matched the direction of intention. Thus, a negative Z obtained under intention to "aim low" was recorded as a positive score. When sufficient data were provided in a report, Z was calculated from those data and compared with the reported results; the new calculation was used if there was a discrepancy. If only probability levels were reported, these were transformed into the corresponding Z score. For experiments reported only as "nonsignificant," a conservative value of $Z = 0$ was assigned; if the outcome was reported only as "statistically significant," $Z = 1.645$ was assigned; and if sample size was not reported or could not be calculated from the information provided, a special code of $N = 1$ was assigned.

3.2. Assessing Quality

Because the hypothesized anomalous effect is not easily accommodated within the prevailing scientific world-view, it is particularly important to assess the trustworthiness of each reviewed experiment. Unfortunately, estimating experimental quality tends to be a subjective task confounded by prior expectations and beliefs.[26,27] Estimates of inter-judge reliability in assessing the quality of research reports, for example, rarely exceed correlations of 0.5.[28] We addressed this problem by assigning to each experiment a single quality weight derived from a set of sixteen binary (present/absent) criteria. The first author coded and double-checked the coding for all studies; the second author independently coded the first 100 studies. Inter-judge reliability for quality criteria was $r = 0.802$ with 98 degrees of freedom.

These criteria were developed from published criticisms about random-number generator experiments[14,15,29-33] and from expert opinion on important methodological considerations when performing studies involving human behavior.[20,34,35] Collectively, these criteria form a measure of credibility by which to judge the reported data. The criteria assess the integrity of the experiment in four categories—procedures,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

statistics, the data, and the RNG device—and they cover virtually all methodological criticisms raised to date. They are (1) control tests noted, (2) local controls conducted, (3) global controls conducted, (4) controls established through the experimental protocol, (5) randomness calibrations conducted, (6) failsafe equipment employed, (7) data automatically recorded, (8) redundant data recording employed, (9) data double checked, (10) data permanently archived, (11) targets alternated on successive trials, (12) data selection prevented by protocol or equipment, (13) fixed run lengths specified, (14) formal experiment declared, (15) tamper-resistant RNG employed, and (16) use of unselected subjects.

Each criterion was coded as being present or absent in the report of an experiment, specifically excluding consideration of previously published descriptions of RNG devices or control tests. This strategy was employed to reflect lower confidence in such experiments since, for example, randomness tests conducted once on an RNG do not guarantee acceptable performance in the same RNG in all future experiments. As a result, assessed quality was conservative, that is, lower than the "true" quality for some experiments, especially those reported only as abstracts or conference proceedings. Using unit weights (which have been shown to be robust in such applications[36]) on each of the sixteen descriptors, the quality rating for an individual experiment was simply the sum of the descriptors. Thus, while a quality score near zero indicated a low quality or poorly reported experiment, a score near sixteen reflected a highly credible experiment.

3.3. Assessing Effect Size

Assume that each of K experiments produces effect size estimates e of a parameter E, based on N samples, and that each e has a known standard error s. The weighted mean effect size is calculated as $e. = \sum \omega_i e_i / \sum \omega_i$, where $\omega_i = 1/s_i^2 = N_i$, and i ranges from 1 to K. The standard error of $e.$ is $s_e = (\sum \omega_i)^{-1/2}$. A test for homogeneity for the K estimates of e_i is given by $H_K = \sum \omega_i (e_i - e.)^2$, where H_K has a chi-square distribution with $K-1$ degrees of freedom.[37] The same procedure can be followed to test for homogeneity of effect size across M independent investigators. In this case, $e._j$ and s_{ej} are calculated per investigator, and the test for homogeneity is performed as $H_M = \sum \omega_j (e._j - e._M)^2$, where $e._j$ and ω_j are mean weighted effect size and $1/s_e^2$ per investigator, respectively, $e._M = \sum \omega_j e._j / \sum \omega_j$, and j ranges from 1 to M. H_M has $M-1$ degrees of freedom.

For a quality-weighted analysis, we may determine $e._Q = \sum (Q_i \omega_i e_i) / \sum (Q_i \omega_i)$, where Q_i is the quality assessed for experiment i. The standard error associated with e_Q is $se_Q = (\sum (Q_i^2 \omega_i) / (\sum Q_i \omega_i)^2)^{-1/2}$; the test for homogeneity is similar to that described above. Finally, following

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the practice of reviewers in the physical sciences,[23,24] we deleted potential "outlier" studies to obtain a homogeneous distribution of effect sizes and to reduce the possibility that the calculated mean effect size may have been spuriously enlarged by extreme values. The procedure used was as follows: If the homogeneity statistic for all studies was significant (at the $p<0.05$ level), the study that would produce the largest reduction in this statistic was deleted; this was repeated until the homogeneity statistic had become nonsignificant.

4. RESULTS

On-line bibliographic databases for psychology and physics journals were searched, as was a specialized database covering parapsychological articles, technical reports, conference proceedings and manuscripts. Altogether 152 references were found from 1959 to 1987. These reports described 832 studies conducted by 68 different investigators (597 experimental studies and 235 control studies). Fifty-four experimental and 33 control studies reported only as nonsignificant were assigned $Z=0$. Six experiments and two control studies coded as $(N=1, Z>0)$ were eliminated from further meta-analysis because effect size could not be accurately estimated (this required the elimination of one investigator who reported a single study). Figures 1 and 2 show the distributions of Z scores reported for control and experimental studies, respectively.

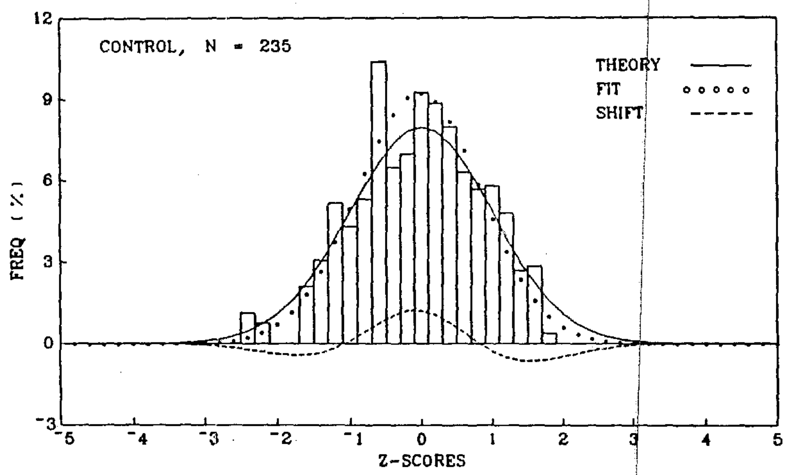


Fig. 1. Distribution of Z scores reported in 235 control studies. Thirty-three of these studies were reported only as "nonsignificant" and were assigned Z scores of zero. To replace the spurious spike at $Z=0$, those 33 studies were recast as normally distributed Z scores, bounded by ± 1.64, averaging $Z=0$.

825 19/12-5

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

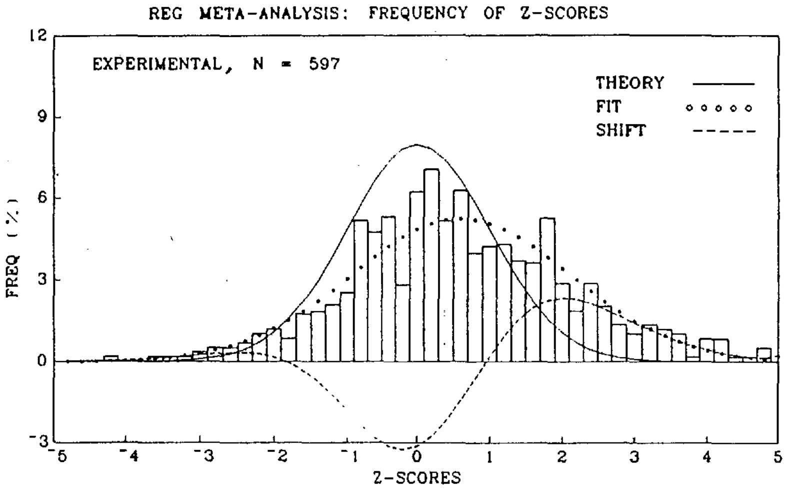


Fig. 2. Distribution of Z scores reported in 597 experimental studies. Fifty-four of these studies were reported as "nonsignificant" and were assigned Z scores of zero. As in Fig. 1, those 54 studies were recast as normally distributed Z scores, bounded by ± 1.64, averaging $Z = 0$.

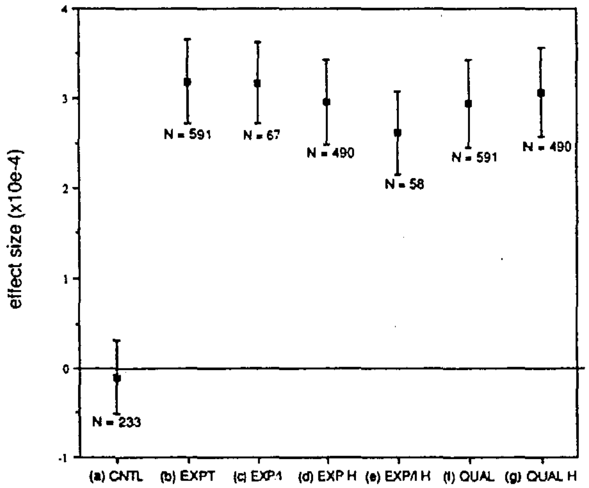


Fig. 3: Mean effect size point estimates ± 1 standard error for (a) control studies and (b) individual experiments; (c) mean effect size per investigator, (d) homogeneous mean effect size for experiments, (e) homogeneous mean effect size per investigator, (f) mean effect size for quality-weighted experiments, and (g) mean effect size for homogeneous quality-weighted experiments.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

These results, expressed as overall mean effect sizes, show that control studies conform well to chance expectation (Fig. 3a), and that experimental effects, whether calculated for studies or investigators, deviate significantly from chance expectation (Fig. 3b, 3c). To obtain a homogeneous distribution of effect sizes, it was necessary to delete 17% of individual outlier studies (Fig. 3d) and 13% of mean effect sizes across investigators (Fig. 3e). This may be compared with exemplary physical and social science reviews, where it is sometimes necessary to discard as many as 45% of the studies to achieve a homogeneous effect size distribution.[19] Of individual studies deleted, 77% deviated from the overall mean in the positive direction, and of investigator means deleted, all were positive (i.e., supportive of the experimental hypothesis).

4.1. Effect of Quality

Some critics have postulated that as experimental quality increases in these studies, effect size would decrease, ultimately regressing to the "true" value of zero, i.e., chance results.[12,13,15,32,33,38] We tested this conjecture with two linear regressions of mean effect size vs. mean quality assessed per investigator, one weighted with ω_j as defined above and the other weighted with the number of studies per investigator. The calculated slope for the former is $-2.5 \times 10^{-5} \pm 3.2 \times 10^{-5}$, and for the latter, $-7.6 \times 10^{-4} \pm 3.9 \times 10^{-4}$. These nonsignificant relationships between quality and effect size is typical of meta-analytic findings in other fields,[39,40] suggesting that the present database is not compromised by poor experimental methodology. Another assessment of the effect of quality was obtained by comparing unweighted and quality-weighted effect sizes per experiment (Fig. 3b vs. 3f). These are nearly identical, and the same is true after deleting outliers to obtain a homogeneous quality-weighted distribution (Fig. 3d vs. 3g), confirming that differences in methodological quality are not significant predictors of effect size.

It might be argued that the quality assessment procedure employed here was nonoptimal because some quality criteria are more important than others, so that if appropriate weights were assigned, the quality-weighted effect size might turn out to be quite different. This was tested by Monte Carlo simulation, using sets of 16 weights, one per criterion, randomly selected over the range 0 to 6. A quality-weighted effect size was calculated for the 597 experiments as before, now using the random weights instead of unit weights, and this process was repeated one thousand times, yielding a distribution of possible quality ratings. The average effect size from the simulation was $3.18 \times 10^{-4} \pm 0.15 \times 10^{-4}$, indicating that in this particular database coded by these sixteen criteria,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the probable range of the quality-weighted mean effect size clearly excludes chance expectation of zero.

4.2. The "Filedrawer" Problem

Although accounting for differences in assessed quality does not nullify the effect, it is well known in the behavioral and social sciences that non-significant studies are published less often than significant studies (this is called the "filedrawer" problem[21,41–43]). If the number of nonsignificant studies in the filedrawer is large, this reporting bias may seriously inflate the effect size estimated in a meta-analysis. We explored several procedures for estimating the magnitude of this problem and to assess the possibility that the filedrawer problem can sufficiently explain the observed results.

The filedrawer hypothesis implicitly maintains that all or nearly all significant positive results are reported. If positive studies are not balanced by reports of studies having chance and negative outcomes, the empirical Z score distribution should show more than the expected proportion of scores in the positive tail beyond $Z = 1.645$. While no argument can be made that all negative effects are reported, it is interesting to note that the database contains 37 Z scores in the negative tail, where only 30 would be expected by chance. On the other hand, there are 152 scores in the positive tail, about five times as many as expected. The question is whether this excess represents a genuine deviation from the null hypothesis or a defect in reporting or editorial practices.

This question may be addressed by modeling based on the assumption that all significant positive results are reported. A four-parameter fit minimizing the chi-square goodness-of-fit statistic was applied to all observed data with $Z \geqslant 1.645$, using the exponential

$$Y = \frac{1}{\sqrt{2}\,\sigma} e^{-\sqrt{2}\,(|x-\mu|/\sigma)} \tag{1}$$

to simulate the effect of skew or kurtosis in producing the disproportionately long positive tail. This exponential is a probability distribution with the same mean and variance as the normal distribution, but with kurtosis = 3.0.

To begin, the null hypothesis of a (0, 1) normal distribution with no kurtosis was considered. To account for the excess in the positive tail, $N = 585{,}000$ filedrawer studies were required, and the chi-squared statistic remained far too large to indicate a reasonable fit (see Table I). This large N, in comparison with the 597 studies actually reported together with the poor goodness-of-fit statistic, suggests that the assumption of a (0, 1) normal distribution is inappropriate.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the probable range of the quality-weighted mean effect size clearly excludes chance expectation of zero.

4.2. The "Filedrawer" Problem

Although accounting for differences in assessed quality does not nullify the effect, it is well known in the behavioral and social sciences that nonsignificant studies are published less often than significant studies (this is called the "filedrawer" problem[21,41-43]). If the number of nonsignificant studies in the filedrawer is large, this reporting bias may seriously inflate the effect size estimated in a meta-analysis. We explored several procedures for estimating the magnitude of this problem and to assess the possibility that the filedrawer problem can sufficiently explain the observed results.

The filedrawer hypothesis implicitly maintains that all or nearly all significant positive results are reported. If positive studies are not balanced by reports of studies having chance and negative outcomes, the empirical Z score distribution should show more than the expected proportion of scores in the positive tail beyond $Z = 1.645$. While no argument can be made that all negative effects are reported, it is interesting to note that the database contains 37 Z scores in the negative tail, where only 30 would be expected by chance. On the other hand, there are 152 scores in the positive tail, about five times as many as expected. The question is whether this excess represents a genuine deviation from the null hypothesis or a defect in reporting or editorial practices.

This question may be addressed by modeling based on the assumption that all significant positive results are reported. A four-parameter fit minimizing the chi-square goodness-of-fit statistic was applied to all observed data with $Z \geqslant 1.645$, using the exponential

$$Y = \frac{1}{\sqrt{2}\,\sigma} e^{-\sqrt{2}(|x-\mu|/\sigma)} \tag{1}$$

to simulate the effect of skew or kurtosis in producing the disproportionately long positive tail. This exponential is a probability distribution with the same mean and variance as the normal distribution, but with kurtosis = 3.0.

To begin, the null hypothesis of a (0, 1) normal distribution with no kurtosis was considered. To account for the excess in the positive tail, $N = 585{,}000$ filedrawer studies were required, and the chi-squared statistic remained far too large to indicate a reasonable fit (see Table I). This large N, in comparison with the 597 studies actually reported together with the poor goodness-of-fit statistic, suggests that the assumption of a (0, 1) normal distribution is inappropriate.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table I. Four-Parameter Fit (E:N, N, Mean, sd) Minimizing Chi-Square (10 df) Goodness-of-Fit Statistic to the Positive Tail of the Observed Z Score Distribution, for Several Exponential:Normal Ratios[a]

Assumption	E:N ratio	N	Mean	sd	Chi-square	p
Normal distribution	0	585,000	0	1	57,867.84	0
(null hypothesis)	1	5,300	0	1	220.97	0
	2	4,800	0	1	167.84	0
	3	4,600	0	1	148.45	0
	10	4,400	0	1	119.69	0
Empirical distribution	0	700	0.145	2.10	23.94	0.008
	1	747	0.345	1.90	16.32	0.091
	2	757	0.445	1.80	14.21	0.164
	3	777	0.445	1.80	11.08	0.226
	10	807	0.445	1.80	11.08	0.351

[a] The null hypothesis is tested by clamping the mean at 0 and the standard deviation at 1, allowing N and E:N to vary. The empirical database is addressed by allowing all four parameters to vary.

Adding simulated kurtosis to a (0, 1) normal distribution by mixing exponential [Eq. (1)] and normal distributions in a 1:1 ratio reduced N by two orders of magnitude, and ratios of 2:1, 3:1, and 10:1 exponential to normal (E:N) yielded further small improvements. However, the chi-squared statistic still indicated a poor fit to the empirical data. Applying the same mixture of exponential and normal distributions, but starting from the observed values of $N = 597$, mean Z score $= 0.645$, and standard deviation $= 1.601$, with the constraint that the mean could only *decrease* from 0.645, resulted in much better fits to the data. Table I shows the results.

This procedure shows that the null hypothesis is unviable, even after allowing a huge filedrawer. The chi-square fit vastly improves with the addition of kurtosis, but only becomes a reasonably good fit when mean and standard deviation are allowed to approximate the empirical values. The filedrawer estimate from this model depends on a number of assumptions (e.g., the true distribution is generally normal, but has a disproportionately large positive tail). It suggests a total number of experimental studies on the order of 800, of which three-fourths have been formally reported.

A somewhat simpler modeling procedure was applied to the data assuming that all studies with significant Z scores in either the positive or negative tail are reported. The model is based on the normal distribution with a standard deviation $= 1$, and estimates the mean and N required to

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

account for the 152 Z scores in the positive tail and 37 Z scores in the negative tail. This mean-shift model, which ignores the shape of the observed distribution, results in an $N = 1,580$ and a mean Z score $= 0.34$.

These modeling efforts suggest that the number of unreported or unretrieved RNG studies falls in the range of 200 to 1,000. A remaining question is, how many filedrawer studies with an average *null* result would be required to reduce the effect to nonsignificance (i.e., $p < 0.05$)? This "failsafe" quantity is 54,000—approximately 90 times the number of studies actually reported. Rosenthal suggests that an effect can be considered robust if the failsafe number is more than five times the observed number of studies.(21)

5. DISCUSSION

Repeatable experiments are the keystone of experimental science. In practice, repeatability depends upon a host of controllable and uncontrollable ingredients, including factors such as stochastic variation, changes in environmental conditions, difficulties in communicating tacit knowledge employed by successful experimenters,(44) and so on. Difficulties in achieving systematic replication are therefore ubiquitous, from experimental psychology(21,45) to particle physics.(23,24) Of course, this is not to say that systematic replication is impossible in these or other fields, but it may appear to be extraordinarily difficult when experiments are considered individually rather than cumulatively. In the case of the present database, the authors of a recent report issued by the US National Research Council stated that the overall results of the RNG experiments could not be explained by chance,(46) but they questioned the quality and replicability of the research. This meta-analysis shows that effects are not a function of experimental quality, and that the replication rate is as good as that found in exemplary experiments in psychology and physics.

Besides the issue of replicability, five other objections are often raised about the present experiments. These are (a) the effect is inconsistent with prevailing scientific models, (b) the experimental methodology is technically naïve, thus the results are not trustworthy, (c) the experiments are vulnerable to fraud by subjects or by experimenters, (d) skeptics cannot obtain positive results, and (e) there are no adequate theoretical explanations or predictions for the anomalous effect.

These criticisms may be addressed as follows: (a) "Inconsistency with the scientific world-view" is essentially a philosophical argument that carries little weight in the face of repeatable experimental evidence, as suggested by the present and two corroborating meta-analyses.(17,18)

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Indeed, if the "inconsistency" argument were sufficient to discount anomalous findings, we would have ignored much of the motivation leading to the development of quantum mechanics. (b) The "naïve methodology" argument was empirically addressed by the assessment of methodological quality in the present analysis. No significant relationship between quality and effect size was found. (c) Fraud postulated as the explanation of the results is untenable as it would have required widespread collusion among 68 independent investigators. In any case, even severe critics of parapsychological experiments have discounted fraud as a viable explanation.[32] (d) Skeptics often assert that only "believers" obtain positive results in such experiments. However, a thorough literature search finds not a single attempted replication of the RNG experiment by a publicly proclaimed skeptic; thus the assertion is not based on verifiable evidence. Furthermore, skeptics who claim to have attempted replications insist (without providing details or references) that they have never achieved positive results in any of their RNG experiments.[15,47] Such a claim is itself quite remarkable, as the likelihood of never obtaining a statistically significant result by chance in series of experiments can be extremely low, depending on the number of experiments conducted. Unfortunately, because we cannot determine how many experiments skeptics have actually conducted, it is impossible to judge the validity of this criticism.

Finally, (e) the "no theoretical basis" argument is correct, but it does not support a negative conclusion about experimental observation. There are at present no adequate theories, with the possible exception of some interpretations of quantum mechanics,[2,3,8,11] that convincingly explain or predict consciousness-related anomalies in random physical systems. We note, however, that the anomalous effects reviewed in this paper apparently can be operationally predicted under well-specified conditions. For example, when individuals are instructed to "aim" for high (or low) numbers in RNG experiments, it is possible to predict with some small degree of confidence that anomalous positive (or negative) shifts of distribution means will be observed.

6. CONCLUSION

In this paper, we have summarized results of all known experiments testing possible interactions between consciousness and the statistical behavior of random-number generators. The overall effect size obtained in experimental conditions cannot be adequately explained by methodological flaws or selective reporting practices. Therefore, after considering all of the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

retrievable evidence, published and unpublished, tempered by all legitimate criticisms raised to date, it is difficult to avoid the conclusion that under certain circumstances, consciousness interacts with random physical systems. Whether this effect will ultimately be established as an overlooked methodological artifact, as a novel bioelectrical perturbation of sensitive electronic devices, or as an empirical contribution to the philosophy of mind, remains to be seen.

ACKNOWLEDGMENTS

This study was supported by major grants from the James S. McDonnell Foundation, Inc. and the John E. Fetzer Foundation, Inc. The authors express their gratitude to Dr. York Dobyns of the Princeton University Engineering Anomalies Laboratory for his assistance with the filedrawer models.

REFERENCES

1. R. G. Jahn and B. J. Dunne, *Margins of Reality* (Harcourt Brace Jovanovich, Orlando, Florida, 1987).
2. B. d'Espagnat, "The quantum theory and reality," *Sci. Am.*, pp. 158–181 (November, 1979).
3. O. Costa de Beauregard, "S-matrix, Feynman zigzag and Einstein correlation," *Phys. Lett.* **67A**, 171–173 (1978).
4. N. D. Mermin, "Is the moon there when nobody looks? Reality and the quantum theory," *Phys. Today*, pp. 38–47 (April, 1985).
5. A. Shimony, "Role of the observer in quantum theory," *Am. J. Phys.* **31**, 755 (1963).
6. E. P. Wigner, "The problem of measurement," *Am. J. Phys.* **31**, 6 (1963).
7. U. Ziemelis, "Quantum-mechanical reality, consciousness and creativity," *Can. Res.* **19**, 62–68 (September, 1986).
8. E. J. Squires, "Many views of one world—an interpretation of quantum theory," *Eur. J. Phys.* **8**, 173 (1987).
9. J. Hall, C. Kim, B. McElroy, and A. Shimony, "Wave-packet reduction as a medium of communication," *Found. Phys.* **7**, 759–767 (1977); p. 761.
10. R. Smith, unpublished manuscript, MIT, 1968. (Cited in Ref. 9, p. 767.)
11. R. G. Jahn and B. J. Dunne, "On the quantum mechanics of consciousness, with application to anomalous phenomena," *Found. Phys.* **16**, 721–772 (1986).
12. J. E. Alcock, *Parapsychology: Science or Magic?* (Pergamon Press, Elmsford, New York, 1981), pp. 124–125.
13. M. Gardner, *Science: Good, Bad, and Bogus* (Prometheus Books, Buffalo, New York, 1981).
14. R. Hyman, "Parapsychological research: A tutorial review and critical appraisal," *Proc. IEEE* **74**, 823–849 (1986).
15. P. Kurtz, "Is parapsychology a science?" In *Paranormal Borderlands of Science*, K. Frazier, ed. (Prometheus Books, Buffalo, New York, 1981).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

16. D. F. Marks, "Investigating the paranormal," *Nature (London)* **320**, 119–124 (1986).
17. C. Honorton, "Replicability, experimenter influence, and parapsychology: An empirical context for the study of mind," paper presented at the annual meeting of the AAAS, Washington, D.C., 1978.
18. E. C. May, B. S. Humphrey, and G. S. Hubbard, "Electronic system perturbation techniques." SRI International Final Report, September 30, 1980.
19. H. Schmidt, "Precognition of a quantum process," *J. Parapsychol.* **33**, 99–108 (1969); "A PK test with electronic equipment," *J. Parapsychol.* **34**, 175–181 (1970); "Mental influence on random events," *New Sci. Sci. J.* **50**, 757–758 (1971); "PK tests with pre-recorded and pre-inspected seed numbers," *J. Parapsychol.* **45**, 87–98 (1981).
20. R. G. Jahn, "The persistent paradox of psychic phenomena: An engineering perspective," *Proc. IEEE* **70**, 136–170 (1982); R. D. Nelson, B. J. Dunne, and R. G. Jahn, "An REG experiment with large data-base capability, III: Operator-related anomalies," Technical Note PEAR 84003, Princeton Engineering Anomalies Research Laboratory, Princeton University, School of Engineering/Applied Science, September 1984; H. Schmidt, R. Morris, and L. Rudolph, "Channeling evidence for a PK effect to independent observers," *J. Parapsychol.* **50**, 1–16 (1986).
21. R. Rosenthal, *Meta-Analytic Procedures for Social Research* (Sage Publications, Beverly Hills, California, 1984); K. Wachter, "Disturbed by meta-analysis?" *Science* **241**, 1407–1408 (1988). We may note that Cohen's h, the difference between control and experimental proportions, is a common effect size measure that might have been used in the present study. This was rejected in favor of e, as defined, because some of the reviewed studies reported only final p values or only overall Z scores; e was thus deemed more useful in the present meta-analysis.
22. R. L. Bangert-Drowns, "Review of developments in meta-analytic method," *Psychol. Bull.* **99**, 388–399 (1986).
23. A. H. Rosenfeld, "The particle data group: Growth and operations." *Annu. Rev. Nucl. Sci.* **25**, 555–599 (1975).
24. C. G. Wohl *et al.*, *Rev. Mod. Phys.* **56**, Part II, p. S5 (1984).
25. G. V. Glass, "In defense of generalization," *Behav. Brain Sci.* **3**, 394–395 (1978).
26. H. M. Cooper, "Scientific guidelines for conducting integrative reviews," *Rev. Educ. Res.* **52**, 291–302 (1982).
27. R. M. Dawes, "You can't systematize human judgment: Dyslexia." In *New Directions for Methodology of Social and Behavioral Science: Fallible Judgment in Behavorial Research*, R. A. Shweder, ed. (Jossey-Bass, San Francisco, 1980), pp. 67–78.
28. S. D. Gottfredson, "Evaluating psychological research reports: Dimensions, reliability, and correlates of quality judgments," *Am. Psychol.* **33**, 920–934 (1978).
29. C. Akers, "Methodological criticisms of parapsychology." In *Advances in Parapsychological Research*, Vol. 4, S. Krippner, ed. (McFarland, Jefferson, North Carolina, 1984); "Can meta-analysis resolve the ESP controversy?" In *A Skeptic's Handbook of Parapsychology*, P. Kurtz, ed. (Prometheus Books, Buffalo, New York, 1985).
30. J. E. Alcock, "Parapsychology: Science of the anomalous or search for the soul," *Behav. Brain Sci.* **10**, 553–565 (1987).
31. P. Diaconis, "Statistical problems in ESP research," *Science* **201**, 131–136 (1978).
32. C. E. M. Hansel, *ESP and Parapsychology: A Critical Reevaluation* (Prometheus Books, Buffalo, New York, 1980).
33. R. Hyman, "The ganzfeld psi experiment: A critical appraisal," *J. Parapsychol.* **49**, 3–50 (1985).
34. T. X. Barber, *Pitfalls in Human Research: Ten Pivotal Points* (Pergamon Press, Elmsford, New York, 1976).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

35. J. B. Rhine, "Comments: 'A new case of experimenter unreliability,'" *J. Parapsychol.* **38**, 215-255 (1974).
36. R. M. Dawes, "The robust beauty of improper linear models in decision making," *Am. Psychol.* **34**, 571-582 (1979).
37. L. V. Hedges, "How hard is hard science, how soft is soft science?" *Am. Psychol.* **42**, 443-455 (1987).
38. C. E. M. Hansel, *ESP: A Scientific Evaluation* (Charles Scribner's Sons, New York, 1966), p. 234.
39. R. Rosenthal and D. B. Rubin, "Interpersonal expectancy effects: The first 345 studies," *Behav. Brain Sci.* **3**, 377-415 (1978).
40. G. V. Glass, B. McGaw, and M. L. Smith, *Meta-analysis in Social Research* (Sage Publications, Beverly Hills, California, 1981).
41. Q. McNemar, "At random: Sense and nonsense," *Am. Psychol.* **15**, 295-300 (1960).
42. S. Iyengar and J. B. Greenhouse, "Selection models and the file-drawer problem," Technical Report 394, Department of Statistics, Carnegie-Mellon University (July, 1987).
43. L. V. Hedges, "Estimation of effect size under nonrandom sampling: The effects of censoring studies yielding statistically insignificant mean differences," *J. Educ. Stat.* **9**, 61-86 (1984).
44. H. H. Collins, *Changing Order: Replication and Induction in Scientific Practice* (Sage Publications, Beverly Hills, California, 1985).
45. S. Epstein, "The stability of behavior, II: Implications for psychological research," *Am. Psychol.* **35**, 790-806 (1980).
46. D. Druckman and J. A. Swets, eds. *Enhancing Human Performance: Issues, Theories, and Techniques* (National Academy Press, Washington, D.C., 1988), p. 207.
47. A. Neher, *The Psychology of Transcendence* (Prentice-Hall, Englewood Cliffs, New Jersey, 1980), p. 147.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Reprinted from the Proceedings of the Parapsychological Association 33rd Annual Convention, Chevy Chase, MD, August 16-20, 1990.

Extraversion and ESP Performance: A Meta-Analysis and a New Confirmation

Charles Honorton and Diane C. Ferrari
Psychophysical Research Laboratories

and

Daryl J. Bem
Cornell University

ABSTRACT— We report a meta-analysis of research on the relationship between performance in extrasensory perception (ESP) tasks and the psychological trait extraversion. The meta-analysis comprises 60 independent studies, 17 independent investigators, and 2,963 subjects. The overall weighted mean correlation is small ($r = .09$), significant ($z = 4.63$, $p = .000004$), and nonhomogeneous. For forced-choice ESP studies ($N = 45$), the ESP/extraversion relationship appears to be an artifact of subjects' knowledge of their ESP performance upon their responses to the extraversion measure: evidence for the relationship is limited to studies where subjects completed the ESP task prior to extraversion assessment ($N = 18$ studies, $r = .17$, $z = 3.51$); no evidence for an ESP/extraversion relationship exists in studies where extraversion was assessed before the ESP task ($N = 16$ studies, $r = -.02$, $z = -0.78$). The two correlations differ significantly ($z = 3.58$, $p = .00045$). For free-response studies, a significant ESP/extraversion relationship exists that is free of this problem: extraversion testing preceded the ESP task in 11 of the 14 free-response studies ($r = .21$, $z = 4.57$, $p = .000005$). The ESP/extraversion relationship is both significant ($r = .20$, $z = 4.46$, $p = .0000083$) and homogeneous in the subset of free-response studies involving individual testing ($N = 12$ studies). The effect is homogeneous across investigators and extraversion scales. We also report a new confirmation of the ESP/extraversion relationship using the Extraversion/Introversion Scale of the Myers-Briggs Type Indicator. The effect size ($r = .18$, $t = 2.67$, 219 *df*, $p = .008$) is very close to the meta-analytic estimate for free-response studies ($r = .20$) and is homogeneous across the eight experimenters. While the relationship between extraversion and ESP in the forced-choice studies is probably artifactual, we conclude that there is a significant ESP/extraversion relationship in the free-response studies, that the relationship is consistent across investigators and scales, and that meta-analysis of parapsychological research domains has predictive validity.

THE RELATIONSHIP BETWEEN ESP performance and individual differences in psychological traits has been explored in many studies since the 1940s. Extraversion is one of the most frequently studied trait variables and three narrative reviews of the ESP/extraversion literature have concluded that ESP performance is positively related to extraversion (Eysenck, 1967; Palmer, 1977; Sargent, 1981). We present a meta-analysis of the extraversion/ESP literature. The purpose of the meta-analysis is to (1) update earlier narrative reviews of the ESP/extraversion relationship, (2) estimate the magnitude of the relationship, (3) assess potential threats to validity, and (4) identify procedural and other variables that moderate the relationship. We will then present a new confirmation of the ESP/extraversion relationship and compare its magnitude to that estimated from the meta-analysis.

The Meta-Analysis

Previous parapsychological meta-analyses have focused on evidence for psi functioning in such research domains as ganzfeld communication (Honorton, 1985), precognition (Honorton & Ferrari, 1989), and studies of the impact of conscious intention on random number generators (Radin & Nelson, 1989). In such cases the effect size index is based on the proportion of hits and the unit of analysis is the trial. In the present case, we are interested in the relationship between psi performance and a predictor variable, extraversion, rather than overall psi perform-

This work was supported by SRI International and by the John E. Fetzer Foundation. We wish to express our appreciation to Edwin C. May of SRI International, who recognized the importance of this work and helped make it possible. We are also grateful to Robert L. Morris, John Palmer and K. Ramakrishna Rao for providing additional information concerning several of the studies used in the meta-analysis, and to H. Kanthamani for valuable comments during the initial phase of this study.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

ance. The effect size index is the correlation coefficient between the two variables and the subject is the unit of analysis.

Method

Retrieval of Studies

We attempted to retrieve all English-language studies of the relationship between extraversion and performance in experimental ESP tasks. The source of studies includes the bibliographies of three narrative reviews (Eysenck, 1967; Palmer, 1977; Sargent, 1981) and inspection of the principal English-language outlets for publication of parapsychological research, including the *Journal of Parapsychology*, *Journal of the American Society for Psychical Research*, *Journal of the Society for Psychical Research*, and *Research in Parapsychology*. In addition, we conducted a computer search of *Psychological Abstracts* using the keywords "extraversion," "extroversion," "introversion," "intraversion," "sociability," and "outgoing."

Investigator Definition

For the purpose of the meta-analysis, we defined independent investigators as investigators who have not worked with other investigators in the data base. In cases of multiple authorship, studies are identified by the senior author of the earliest publication in the data base. For example, studies by Kanthamani (1966), Kanthamani and Rao (1972) and Krishna and Rao (1981) are all indentified as Kanthamani studies. While Kanthamani and Krishna did not work together, they share Rao as a coauthor and are therefore considered to represent a single investigator set. Similarly, studies by Humphrey (1945, 1951) and by Nicol and Humphrey (1953, 1955) are all indentified as Humphrey studies. Laboratory affiliation was not used to identify independent investigators since several investigators worked in different laboratories and several generations of investigators worked in one laboratory.

Procedural Features

Besides bibliographical data identifying the investigator, publication source and date, we coded various procedural, sampling, and statistical features for each study. The procedural features include the type of ESP task (forced-choice or free-response), test setting (individual or group testing), ESP mode (telepathy, clairvoyance, precognition, mixed), and amount of feedback in the ESP task. Procedural features relevant to assessing research quality are described in the following section. Sampling and statistical features coded include the number of subjects, the subject population used, the instrument used for measuring extraversion, and the inferential statistics reported for testing the relationship between ESP performance and extraversion.

Criteria for Assessing Research Quality

Methodological variables were coded in terms of procedural descriptions (or their absence) in the research reports. This approach was used in an earlier meta-analysis of the ESP ganzfeld domain (Honorton, 1985), resulting in study quality ratings that were generally in agreement ($r_{26} = .77$, $p = 10^{-6}$) with independent "flaw" ratings by an outside critic (Hyman, 1985). Two sets of criteria were used. One set assessed threats to the validity of the ESP measure. The other set assessed threats to the validity of the relationship between the ESP and extraversion measures.

The ESP quality analysis includes four criteria. One point was given (or withheld) for each of the following:

Control against sensory leakage. Credit was given to GESP (telepathy) studies if the report specified that each of the following criteria were met: sender and receiver were located in separate rooms, the test situation prohibited auditory or other cues from sender to receiver, the sender and receiver were monitored by experimenters, and in free-response studies involving subject judging, duplicate target sets were employed. Credit was given to clairvoyance studies if the report specified that cues from the targets were prohibited by means of physical distance, screens, or opaque packaging. Precognition studies were considered to be immune to sensory leakage problems.

Randomization. Studies received credit if random number tables, random number generators, or mechanical shufflers were used to randomize the targets. Studies using informal methods of randomization such as hand shuffling and dice throwing, or in which there was no randomization, received no credit. Reports failing to identify the method of randomization received no credit.

Duplicate recording. Studies reporting duplicate recording of targets and responses received credit. Reports that failed to describe their data recording procedures and those in which data recording was performed by a single experimenter, received no credit.

Duplicate scoring. Studies reporting duplicate checking of hits received credit. Reports that failed to describe their scoring procedures and those in which scoring was performed by a single experimenter, received no credit.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Figure 1. Stem-and-leaf frequency distribution of correlation coefficients between extraversion and ESP performance (N = 47).

-.4	4
-.3	6
-.2	5 2
-.1	7 4 1
-.0	9 2 0 0
.0	0 5 8 9
.1	0 0 1 2 2 7 8 8
.2	0 3 3 8
.3	0 0 4 4 4 5 7 8 8
.4	0 6 6 9 9
.5	3 4 8
.6	6 7
.7	
.8	
.9	1

We assessed two aspects of methodology that could result in a spurious correlation between ESP scores and measures of extraversion: advance specification of the criteria used to define extraversion and the order in which the ESP task and extraversion measure were administered.

A priori definition of extraversion. An inflated relationship between ESP and extraversion scores could occur if the investigator selected an "optimal" extraversion/introversion breakdown after observing the data, without correcting for multiple analysis. We coded studies as to whether the classification of extraversion/introversion appeared to be predetermined or post hoc. No credit was given for studies using nonstandard classifications (e.g., 1 *sd*, median split, etc.) unless the report explicitly stated that the method of classification was preplanned.

Order of administration of extraversion and ESP measures. A spurious correlation between ESP scores and extraversion could arise if subjects' responses on the extraversion scale were influenced by knowledge of their performance in the ESP task. There is some evidence that subjects' responses to psychological tests may be influenced by feedback concerning their ESP performance (Palmer & Lieberman, 1975). We coded studies as to whether the extraversion scale was given before or after the ESP test. This information was available for 45 of the 60 studies.

Meta-Analysis of Correlation Coefficients

We combined correlations across independent studies using the procedures described by Hedges and Olkin (1985) and Rosenthal (1984). All statistics were converted to indices of association; *t* tests were converted to point-biserials and phi coefficients were computed from 2 x 2 contingency tables.[1] We estimated unreported correlations from the reported *p* values (with results reported only as "nonsignificant" obtained in 13 cases set equal to .00). The signs of the correlations were adjusted if necessary to insure that positive correlations reflect positive relationships between extraversion and ESP performance. The correlations were pooled across studies but within categories based on methodological features associated with the studies. The correlations were transformed to their Fisher's z equivalents, weighted by their *df*, and averaged. We determined the two-tailed significance levels and 95% confidence intervals (CIs; Hedges and Olkin, 1985). Finally, we conducted chi-square tests of homogeneity (Hedges and Olkin, 1985; Rosenthal, 1984) and transformed mean zs back to the *r* metric.

The chi-square homogeneity test assesses the consistency of study outcomes, providing a quantitative index of replicability. A set of studies are exact replicates if their effect sizes are identical. They are homogeneous if the variability of effect sizes can be explained by sampling error. A significant but nonhomogeneous effect indicates the presence of moderating variables and homogeneity tests are used to identify moderating variables by subdividing studies into smaller, methodologically similar subgroups (Hedges, 1987).

Results

We retrieved 60 independent studies contained in 35 publications by 17 independent investigators. The studies were reported over a span of 38 years, between 1945 and 1983. The data base comprises 2,963 subjects. Forty-five studies involve ESP card-guessing tasks or similar forced-choice tasks. Fourteen studies employed free-response ESP tasks, and one involved a remote physiological influence task.

The unweighted correlations range from –0.44 to .91. Figure 1 shows a stem-and-leaf display (Tukey, 1977) of the correlation coefficients. (The 13 studies that were assigned *r*'s of zero because of insufficient information are omitted). Unlike other methods of

[1] Two studies provided only trial-based tests (CRd). Correlations were estimated for these studies using a method for estimating effect sizes from critical ratios reported by McCarthy & Schechter (1986). Their method provides an estimate of Cohen's *d* which we then converted to the *r* metric.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 1. Summary of Extraversion—ESP Meta-Analysis

	Number of independent effect sizes	Number of independent investigators	Total number of subjects	Mean effect size (r)	95% confidence interval From	To	z	p	$\chi^2(k-1)$
Overall, all studies	60	17	2963	.09	.05	.12	4.63	.000004	126.21*
Studies with r available	47	14	1853	.14	.10	.19	5.90	<.000001	115.91*
Forced-choice guessing tasks	45	13	2169	.06	.02	.11	2.86	.0042	92.82*
Individual testing	21	11	920	.15	.09	.22	4.54	.000006	42.99*
Group testing	24	8	1249	–.00	–.06	.05	–0.12	.904	37.35
Free-response imaging tasks[a]	14	4	612	.20	.12	.28	4.82	.0000015	23.40*
Individual testing	12	3	512	.20	.11	.29	4.46	.0000083	15.85
Group testing	2	1	100	.19	–.01	.37	1.83	.067	7.53*

Notes. r is the weighted average correlation coefficient (Hedges & Olkin, 1985). k represents the number of independent effect sizes. χ^2 is the within group homogeneity statistic (Rosenthal, 1984).

displaying frequency distributions, the stem-and-leaf plot retains the numerical data precisely. (Turned on its side, the stem-and-leaf plot is a histogram.) Each number includes a stem and one or more leaves. For example, the stem .2 is followed by leaves of 0, 3, 3, 8 representing r's of .20, .23, .23, and .28.

The meta-analysis is summarized in Table 1. The study grouping is identified in the first column. Columns two through four show, respectively, the number of studies, investigators, and subjects. The mean weighted effect size (r) is shown in column five and columns six and seven show, respectively, the lower and upper 95% (CI) for the effect size. The cumulative z-score and its associated two-tailed p-value are shown in columns eight and nine. The chi-square homogeneity statistic is presented in the last column.

Overall Results

The results for all 60 studies are shown in the first row. The mean weighted r is 0.09 ($z = 4.63$, $p = .000004$, two-tailed). The 95% CI is an r from .05 to .12. While significant, the study effect sizes are nonhomogeneous ($\chi^2_{59} = 126.21, p < .05$). The second row shows the same analysis, omitting the 13 studies that were assigned r's of zero.

Forced-choice Studies

The ESP/extraversion correlations for the 45 forced-choice (FC) studies are presented in row three of Table 1. These studies were contributed by 13 independent investigators and include 2169 subjects. The mean weighted r is .06 ($z = 2.86, p = .0042$, 95% CI from .02 to .11). The FC correlations are significantly nonhomogeneous ($\chi^2_{44} = 92.82, p < .05$).

FC studies involving individual testing. Twenty-one of the FC studies involved individual testing (Table 1, row 4). These studies were performed by 11 independent investigators and included 920 subjects. The mean weighted r is .15 ($z = 4.54, p = .000006$, 95% CI from .09 to .22). The correlations are nonhomogeneous ($\chi^2_{20} = 42.99, p < .05$).

FC studies involving group testing. The remaining 24 FC studies involved group testing. These studies were conducted by eight independent investigators and included 1249 subjects. They yield a mean weighted r of –0.00 ($z = -0.12, p = .904$, 95% CI from –.06 to .05). The group FC studies are homogeneous ($\chi^2_{23} = 37.35, p > .05$). Thus, the FC studies involving group testing yielded uniformly null correlations between ESP performance and extraversion.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Individual versus group testing. Evidence for a relationship between extraversion and forced-choice ESP performance is limited to FC studies involving individual testing. The difference between the ESP/extraversion correlations for individual and group testing is significant (Cohen's $q = .14$, $z = 3.47$, $p = .00052$, two-tailed).

Free-response Studies

The results for the 14 free-response (FR) studies are shown in row six of Table 1. The FR studies were contributed by four independent investigators and included 612 subjects. The mean weighted r is .20 ($z = 4.82$, $p = .0000015$, two-tailed). The 95% CI is an r from .12 to .28. The overall FR outcomes, while highly significant, are significantly nonhomogeneous ($\chi^2_{13} = 23.40$, $p < .05$). This nonhomogeneity is due to a moderating variable, test setting.

Free-response studies involving individual testing. Twelve FR studies employed individual testing (Table 1, row 7). These studies were contributed by three independent investigators and include 512 subjects. The results are both significant and homogeneous. The mean weighted r is .20 ($z = 4.46$, $p = .0000083$, 95% CI from .11 to .29, $\chi^2_{11} = 15.85$ $p > .05$).

Free-response studies involving group testing. Only two FR studies involved group testing (Table 1, row 8). Both studies were contributed by the same investigator. The mean weighted r is .19 ($z = 1.83$, $p = .067$, 95% CI from –.01 to .37). The results are significantly nonhomogeneous ($\chi^2_1 = 7.53$, $p < .05$).

Free-response versus Forced-choice Studies

The mean correlation between ESP performance and extraversion is significantly larger in studies with free-response ESP tests than in those using forced-choice tests (Cohen's $q = .14$, $z = 3.11$, $p = .0019$, two-tailed).

Quality Analysis of Forced-Choice Studies

Sensory leakage. Thirty-one of the FC studies describe the use of methods that satisfy our criteria for adequate control against sensory leakage in the ESP task. The ESP/extraversion relationship is not significant for these studies; the mean weighted r is .04 ($z = 1.47$, $p = .142$, two-tailed). The ESP/extraversion relationship is significant in the 14 FC studies that are amenable to sensory leakage ($r = .14$, $z = 3.20$, $p = .0014$, two-tailed). The difference between the two correlations is significant (Cohen's $q = -.10$, $z = -2.03$, $p = .042$, two-tailed).

Randomization. Thirteen FC studies satisfy our criteria for randomization. The mean weighted r for these studies is .13 ($z = 2.27$, $p = .023$). Thirty-two FC studies failed our randomization criteria; they yielded a mean weighted r of .05 ($z = 2.13$, $p = .033$, two-tailed). The ESP/extraversion correlation is nonsignificantly *lower* in these studies than in studies using formal randomization methods (Cohen's $q = .08$, $z = 1.40$, $p = .162$, two-tailed).

Duplicate recording. Six FC studies employed duplicate recording of targets and responses. The mean weighted r for these studies is .31 ($z = 3.52$, $p = .00043$, two-tailed). The 39 studies without duplicate recording yielded a mean weighted r of .05 ($z = 2.01$, $p = .0444$, two-tailed). The ESP/extraversion relationship is significantly *stronger* in studies with duplicate recording (Cohen's $q = .27$, $z = 3.02$, $p = .0025$).

Duplicate checking. Eighteen FC studies reported using methods for duplicate checking of hits. These studies are associated with a mean weighted r of .27 ($z = 5.72$, $p < 10^{-6}$, two-tailed). The 27 FC studies that did not use methods for duplicate checking yielded a nonsignificant r of .01 ($z = 0.23$, $p = .59$). The difference between the two correlations is significant (Cohen's $q = .27$, $z = 5.20$, $p < 10^{-6}$, two-tailed).

A priori definition of extraversion. Thirty-seven FC studies satisfied our criteria for *a priori* definition of extraversion. These studies had a mean weighted r of .05 ($z = 2.20$, $p = .028$). The remaining eight FC studies failed our criteria for *a priori* definition of extraversion ($r = .10$, $z = 2.00$, $p = .046$, two-tailed). The two correlations do not differ significantly (Cohen's $q = -.05$, $z = 0.92$, $p = .358$, two-tailed).

Order of administration of extraversion and ESP measures. Evidence for a relationship between forced-choice ESP performance and extraversion is entirely dependent upon the outcomes of studies in which extraversion was measured after the ESP test ($N = 18$ studies, $r = .17$, $z = 3.51$, $p = .00045$). The correlation between FC ESP performance and extraversion in studies in which extraversion was measured *before* the ESP task is not significant ($N = 16$ studies, $r = -.02$, $z = -0.78$, $p = .782$) and the difference between the two correlations is significant (Cohen's $q = .19$, $z = 3.58$, $p = .00034$, two-tailed). This difference is not attributable to methodological features of the two groups such as test setting, the extraversion measure, or ESP mode.

Degree of ESP feedback is documented in nine of the 18 studies that measured extraversion after the ESP task and subjects received feedback of their ESP performance in each of these studies ($r = .29$, $z = 4.59$,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 2. Forced-choice Outcomes by Investigator

Investigator	*N* studies	*N* subjects	*r*	*z*	*p*
Astrom – ?	1	48	.24	1.63	.103
Casper – A	1	20	.53	2.46	.014
Green	2	148	.00	0.00	.500
?	1	108	.00	0.00	.500
A	1	40	.00	0.00	.500
Humphrey	6	138	.26	2.86	.0042
?	3	55	.27	1.84	.0658
A	3	83	.22	1.95	.051
Kanthamani	7	301	.21	3.59	.00033
?	1	60	.00	0.00	.500
A	3	108	.38	4.03	.000056
B	3	133	.02	0.25	.400
McElroy – A	1	31	.00	0.00	.500
Nash	8	207	.14	1.92	.054
?	2	60	.29	2.22	.026
A	6	147	.08	0.86	.390
Nielsen	3	60	-.04	-0.29	.771
A	2	48	.00	0.00	.500
B	1	12	-.23	-0.69	.755
Sargent	3	85	.17	1.46	.144
B	2	40	-.02	-0.12	.548
M	1	45	.31	2.07	.038
Shields – ?	2	99	.30	2.98	.0029
Shrager – B	2	76	.18	1.48	.139
Szczygielski – A	1	17	-.36	-1.42	.922
Thalbourne – B	8	939	-.04	-1.28	.90

Overall Forced-choice z by investigators = 3.49, p = .00048, two-tailed, $\chi^2_{12} = 41.20$, $p < .05$

Extraversion measured before ESP test (5 investigators): z by investigators = –0.71, p = .761, two-tailed, $\chi^2_4 = 3.97$, $p > .05$

Extraversion measured after ESP test (8 investigators): z by investigators = 3.51, p = .00045, two-tailed, $\chi^2_7 = 17.29$, $p < .05$

Notes. *r* is the weighted average correlation coefficient (Hedges & Olkin, 1985). The letters B and A following the investigator's name indicate whether the extraversion/introversion measure was administed before (B) or after (A) the ESP task. Studies where this information is not available are indicated by a question mark (?). Testing order was mixed (M) in one study. χ^2 is the within group homogeneity statistic (Rosenthal, 1984).

p = .0000045, two-tailed). Seven of the nine studies where feedback is undocumented were group studies which usually involve delayed feedback or no feedback at all. These studies yield a nonsignificant correlation between performance and extraversion (r = .05, z = 0.64, p = .522, two-tailed) which is significantly lower than that for the studies known to involve feedback (Cohen's q = .25, z = 2.54, p = .011, two-tailed). The relationship between forced-choice ESP performance and extraversion thus appears to be artifactual.

Patterns over time. The forced-choice studies were reported between 1945 and 1982. There is a significant decline in the magnitude of the ESP/extraversion relationship over this period (r = –.40, t = –2.89, 43 df, p = .006, two-tailed). Moreover, methodological quality, as assessed in terms of threats to the validity of the ESP measure, has not improved over the survey period (r = .01, t = 0.03, p = .976, two-tailed). These results are contrary to the patterns found in meta-analyses of three other parapsychological domains, which exhibit constant effect sizes and significant

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

methodological improvement over time (Honorton, 1985; Honorton & Ferrari, 1989; Radin & Nelson, 1989). There has been substantial improvement with regard to threats to the validity of the ESP/extraversion relationship; more recently reported studies have generally involved administration of the extraversion measure before the ESP task ($r = .78$, $t = 7.02$, 32 *df*, $p<10^{-6}$). These findings are consistent with the conclusion that the FC ESP/extraversion relationship is artifactual.

FC Outcomes by Investigator

The FC study outcomes by investigator are shown in Table 2. The order of ESP and extraversion testing is indicated following the investigator's name for investigators with single studies or multiple studies involving uniform testing order. Separate breakdowns are given for investigators with studies involving different testing orders. Significant outcomes were obtained by four of the 13 FC investigators (31%); using the investigator as the unit of analysis, the overall results are significant ($z = 3.49$, $p = .00048$, two-tailed) but nonhomogeneous ($\chi^2_{12} = 41.20$, $p < .05$). The effect of testing order accounts for the overall significance and nonhomogeneity. The five investigators who measured extraversion before the ESP task have outcomes that are nonsignificant and homogeneous ($z = -0.71$, $p = .761$, two-tailed, $\chi^2_4 = 3.97$, $p > .05$), while the outcomes of the eight investigators who measured extraversion after the ESP task are significant and nonhomogeneous ($z = 3.51$, $p = .00045$, two-tailed, $\chi^2_7 = 17.29$, $p < .05$). Thus, the impact of ESP/extraversion testing order is consistent across investigators and is not attributable to idiosyncratic research styles or other characteristics of a single prolific investigator.

Quality Analysis of Free-response Studies

Sensory leakage. All 14 FR studies satisfied our criteria for adequacy of control against sensory leakage.

Randomization. Nine FR studies satisfied our criteria for randomization ($r = .38$, $z = 4.74$, $p = .0000022$, two-tailed). Five FR studies employed informal randomization procedures or failed to document their method of randomization ($r = .14$, $z = 2.86$, $p = .0042$, two-tailed). The difference between the two correlations is significant (Cohen's $q = .26$, $z = 2.86$, $p = .0042$, two-tailed).

Duplicate recording. Thirteen of the 14 FR studies employed duplicate recording methods ($r = .29$, $z = 4.62$, $p = .0000039$, two-tailed).

Duplicate checking. All 14 FR studies employed duplicate checking methods.

A priori definition of extraversion. Eleven FR studies satisfied our criteria for *a priori* definition of extraversion ($r = .16$, $z = 3.72$, $p = .0002$, two-tailed). Three studies failed to document their basis of classification ($r = .48$, $z = 4.01$, $p = .000061$, two-tailed). The two correlations differ significantly (Cohen's $q = .36$, $z = 2.77$, $p = 0.0056$, two-tailed).

Order of administration of extraversion and ESP measures. The extraversion scale was administered before the ESP task in 11 of the FR studies ($r = .21$, $z = 4.57$, $p = .000005$, two-tailed). The remaining three FR studies failed to report the order in which the ESP and extraversion measures were given ($r = .15$, $z = 1.64$, $p = .101$, two-tailed). The difference between the two correlations is not significant (Cohen's $q = .06$, $z = 0.62$, $p = .532$). Thus, for the free-response studies, the evidence for a relationship between ESP performance and extraversion is not susceptible to explanation in terms of an order artifact.

Patterns over time. The free-response studies were reported between 1960 and 1982. Unlike the forced-choice studies, the magnitude of the ESP/extraversion relationship has increased over time, though not significantly so ($r = .19$, $t = 0.66$, 12 *df*, $p = .524$, two-tailed). The methodological quality of the free-response studies has also improved over time ($r = .36$, $t = 1.35$, 12 *df*, $p = .202$, two-tailed).

Confirmation of Differences in Research Quality in relation to Test Setting

Honorton & Ferrari (1989), in a meta-analysis of forced-choice precognition experiments, found that studies involving individual testing were of significantly higher methodological quality than studies involving group testing ($t = 3.08$, 137 *df*, $p = .003$, two-tailed). We have confirmed this finding in the ESP/extraversion meta-analysis ($t = 2.27$, 39 *df*, $p = .015$, one-tailed). This analysis excludes the 12 precognition studies which overlap with the earlier meta-analysis.

Since we have determined that there is no ESP/extraversion relationship in the forced-choice studies when the effects of task order are considered, the remaining analyses are restricted to the free-response studies.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 3. Free-response Outcomes by Investigator

Investigator	N studies	N subjects	r	z	p
Bellis & Morris	1	23	.47	2.26	.024
Braud	2	100	.19	1.86	.063
Marsh	1	311	.13	2.30	.021
Sargent	10	178	.31	3.85	.00012

z by investigators = 5.11, $p = 3 \times 10^{-7}$, two-tailed

$\chi^2_3 = 2.51, p > .05$

Notes. r is the weighted average correlation coefficient (Hedges & Olkin, 1985). χ^2 is the within group homogeneity statistic (Rosenthal, 1984).

Consistency across Investigators

Table 3 shows the overall FR results by investigator. Three of the four investigators have significant ESP/extraversion correlations and the results of the fourth investigator (Braud) approach significance. The z by investigator is 5.11, a result that should arise by chance less than one time in 3.3 million. The results are homogeneous across investigators ($\chi^2_3 = 2.51$, $p > .05$). Although 10 of the 14 FR studies were contributed by one investigator (Sargent), evidence for the relationship between free-response ESP performance and extraversion is not dependent upon that investigator. When Sargent's work is eliminated, the results of the three remaining investigators still strongly supports a relationship between ESP performance and extraversion ($z = 3.35$, $p = 0.0008$, two-tailed). Therefore, we conclude that the ESP/extraversion relationship is consistent across investigators.

Extraversion Measures

Each FR investigator used a different scale for measuring extraversion. Marsh used the Bernreuter Personality Inventory (Super, 1942); Sargent and his group used the Cattell 16PF (Cattell, Eber & Tatsuoka, 1970); Braud and Bellis & Morris used scales constructed by the investigators (with no psychometric validation provided). It is impossible to isolate the effects of the instruments for measuring extraversion from the ensemble of procedures and research styles associated with the investigators. All that can be said is that a relationship between extraversion and ESP performance is evident in studies using four different measures of extraversion.

Selective Reporting

In order to assess the vulnerability of these studies to selective reporting, we used Rosenthal's (1984) "Fail-safe N" statistic to estimate the number of unreported studies averaging null outcomes necessary to reduce the known data base to nonsigificance. The Fail-safe N is 140 studies. In other words, if we were to assume that the observed outcomes arise from selective reporting, it would be necessary to postulate 10 unreported studies averaging null outcomes for each reported study. Therefore, we conclude that the free-response ESP/extraversion relationship cannot be explained on the basis of selective reporting.

Power Analysis

The FR mean r of .20 is equivalent to an average ESP scoring advantage for extraverts over introverts of 0.4 standard deviations. The FR studies average sample size is 44 subjects and the likelihood of detecting a correlation of .2 at the five percent significance level with this sample size—the statistical power—is 37 percent (Cohen, 1977, p. 87). Thus, in a sample of 14 studies, the expected number of statistically significant studies is 5.2; the actual number of significant studies is seven (exact binomial probability, with $p = .37$ & $q = .63$, = .23, one-tailed). Thus, the observed rate of significant outcomes is consistent with a correlation of .2.

Achievement of statistical significance, assuming a correlation of .2, is essentially a coin toss with sample sizes less than 68 subjects; a sample size of 180 is necessary to achieve 85 percent power.

In the following section, we explore the predictive validity of the ESP/extraversion meta-analysis by comparing the meta-analytic estimate to the outcome of a new data set.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 4. ESP/Extraversion Correlations by Experimenter in the PRL Novice Series

Experimenter	*N* Subjects	*r*	z	Experimenter EI Score
Honorton	41	.27	1.71	101
Quant	69	.29	2.38	103
Derr	22	.03	0.68	81
Berger	13	–.37	–1.18	115
Varvoglis	21	.08	0.32	133
Schechter	7	–.05	–0.10	125
Ferrari	10	–.20	–0.54	133
Schlitz	7	.15	0.92	69

Note. *r* is the weighted average correlation coefficient (Hedges & Olkin, 1985).

A New Confirmation

Extraversion data is available for 221 of the 241 subjects in a series of ESP ganzfeld studies reported by Honorton, Berger, Varvoglis, Quant, Derr, Hansen, Schechter & Ferrari (1990) and conducted at the Psychophysical Research Laboratories (PRL) in Princeton, N.J. The experimental procedures are described in detail in the Honorton, et al. (1990) report.

Subjects

The subjects were 131 women and 90 men. Their average age is 37 years (*sd* = 11.7). This is a well-educated group; the mean formal education is 15.5 years (*sd* = 2.0) and belief in psi is strong in this population. On a seven-point scale where "1" indicates strong disbelief and "7" indicates strong belief in psi, the mean is 6.20 (*sd* = 1.03). Personal experiences suggestive of psi were reported by 88 percent of the subjects; eighty percent reported ostensible telepathic experiences. Eighty percent have had some training in meditation or other techniques involving internal focus of attention. One hundred and sixty-three subjects contributed a single ESP ganzfeld session and 58 contributed multiple sessions.

Extraversion Measure

Extraversion was measured using the continuous scores of the Extraversion/Introversion (EI) Scale in Form F of the Myers-Briggs Type Indicator (MBTI; Briggs & Myers, 1957). The MBTI was not used in any of the meta-analysis studies. The MBTI EI Scale is constructed so that scores below 100 indicate extraversion and scores above 100 indicate introversion. (For consistency with the meta-analysis, we have reversed the signs so that positive correlations reflect a positive relationship between ESP performance and extraversion.) The mean EI score for the PRL subjects is 100.36 (*sd* = 25.18).

ESP Measure

ESP performance was measured using the standardized ratings of the target and decoys (Stanford's z-scores; Stanford and Sargent, 1983). Stanford *z*'s were averaged for subjects with multiple sessions.

Results

Overall results. The correlation between ESP performance and extraversion in the PRL series is significant ($r = .18$, 219 *df*, $t = 2.67$, $p = .008$, two-tailed, 95% CI from .05 to .30). This outcome is very close to the meta-analytic estimate for free-response studies ($r = .20$) and the difference between the two correlations is nonsignificant (Cohen's $q = .02$, $z = -0.26$, $p = .793$, two-tailed).

Ganzfeld Novices. The results are similar if we restrict our analysis to the five PRL Novice series with inexperienced subjects who each completed a single ganzfeld session. MBTI data is available for 190 of the 205 Novices and the mean weighted *r* for the five series is .17 ($z = 2.25$, $p = .024$, two-tailed, 95% CI from .02 to .31). The ESP/extraversion correlations are homogeneous across the five series ($\chi^2_4 = 2.88$, $p > .05$). Eleven subjects in the first Novice series (Series 101) completed the MBTI between six and eighteen months after their ESP ganzfeld session and we did not maintain records of their identity. However, the results are essentially the same when this series is eliminated. The mean weighted *r* for the remaining four Novice series is .19 ($z = 2.30$, $p = .021$, two-tailed, 95% CI from .03 to .34).

Outcome by experimenter. Eight experimenters contributed to the PRL data base (Honorton, et al., 1990). Table 4 shows the ESP/extraversion correlation by

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

experimenter for the five Novice series. The mean weighted r for the eight experimenters is .16 ($z = 2.09$, $p = .037$, two-tailed, 95% CI from .01 to .30). The results are homogeneous across the eight experimenters ($\chi^2_7 = 6.43, p > .05$).

Outcome in relation to EI status of experimenter. It is possible that the relationship between ESP performance and extraversion is moderated by personality characteristics of the experimenter. The last column of Table 4 shows the MBTI EI scores for each experimenter. Only two experimenters (Derr and Schlitz) are extraverts. Two others (Honorton and Quant) are borderline introverts. While the above analyses indicate that the ESP/extraversion correlation is consistent across experimenters, there is a nonsignificant tendency for the relationship to be stronger in the data of less introverted experimenters ($r = .47$, 6 df, $p = .235$, two-tailed).

Combined Estimate of the Relationship between Free-response ESP Performance and Extraversion

Combining the new confirmation with the meta-analysis, the overall mean weighted r is .19 ($z = 5.50$, $p = 3.8 \times 10^{-8}$, 95% CI from .13 to .26). The "Fail-safe N" for the combined estimate is 181 studies, or a ratio of 12 unreported studies averaging null effects for each known study. Four of the five investigators have overall significant outcomes and the outcomes are homogeneous across investigators $\chi^2_4 = 6.03$, $p > .05$).

Discussion

The Meta-Analysis

Forced-choice studies. The meta-analysis challenges the conclusions from earlier narrative reviews of the relationship between extraversion and forced-choice ESP performance (Eysenck, 1967; Palmer, 1977; Sargent, 1981). The apparent relationship between extraversion and ESP performance in these studies appears to be due to the influence of subjects' knowledge of their ESP performance on their subsequent responses to the extraversion measures. Evidence for a relationship between ESP and extraversion occurs only when extraversion was measured after the ESP test; no evidence of an ESP/extraversion relationship is found in studies where extraversion was measured before the ESP task.

Evidence for a nonzero effect in the forced-choice studies is also limited to the subset of studies involving ESP testing procedures that were vulnerable to potential sensory leakage. There is reason to believe, however, that this may result from a procedural confound: six the eight studies in this subgroup for which information on the order of testing is available also involved extraversion testing following ESP feedback.

The apparent biasing effect of ESP feedback probably arises from one of two possibilities. Awareness of "success" or "failure" may lead subjects to later perceive themselves as more extraverted or introverted. Or, the problem may arise from an experimenter expectancy effect (Rosenthal & Rubin, 1978), in which subjects respond to the investigator's expectations that extraverts are more successful in ESP tasks than introverts. Obviously, further research will be necessary to clarify the problem.

The existence of this problem, however, necessarily arouses concern over the viability of reported relationships between ESP performance and other personality factors such as neuroticism (Palmer, 1977). Much of the research in these areas was conducted by the same investigators, and it is likely that similar methods were used. We believe that conclusions regarding the relationship between ESP performance and other personality variables should be suspended until the relevant study domains can be examined with respect to this problem.

Free-response studies. The meta-analysis does support the existence of a relationship between extraversion and free-response ESP performance. The free-response studies are not amenable to explanation in terms of an order artifact or other identifiable threats to validity. The overall correlation of .20 would be expected to occur only about one time in 674,000 by chance. Three of the four investigators contributing to this data base obtained significant ESP/extraversion relationships, and the fourth investigator's results approach significance. The correlations are homogeneous across investigators, and across the largest grouping of studies in which subjects were tested individually. The effect remains highly significant even when 71 percent of the studies, contributed by one investigator, are eliminated from consideration. Thus, the relationship seems to be robust. Estimation of the filedrawer problem (Rosenthal, 1984), indicates that it would be necessary to postulate 10 unreported studies averaging null results for every retrieved study in order to account for the observed effect on the basis of selective reporting.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The New Confirmation

The results of the confirmation, involving a new set of investigators and a new scale of extraversion, support the meta-analytic findings and increase their generalizability. The relationship between free-response ESP performance and extraversion now spans 833 subjects and five independent investigator teams. The homogeneity of the effect across the eight experimenters in the confirmatory study further increases our confidence that the effect is replicable and is not dependent upon unknown characteristics of individual investigators. A nonsignificant trend in the data does suggest that the ESP/extraversion relationship may, to some extent, be moderated by the experimenter's extravertedness and it may be advisable for future investigators to record and report extraversion/introversion scores of the experimenters.

The Predictive Validity of Meta-Analysis

Meta-analysis is a powerful tool for summarizing existing evidence. It enables more precise estimation of the significance and magnitude of behavioral effects than has been possible with traditional narrative reviews, and is useful in identifying moderating variables. In the present case, meta-analytic techniques revealed a serious source of bias that had been overlooked in earlier narrative reviews of the ESP/extraversion domain. Moreover, the meta-analysis identified a subset of the domain that is not amenable to the discovered bias and provided an estimate of the magnitude of the relationship between ESP and extraversion in that subset.

Ultimately, the usefulness of meta-analysis will be judged by its ability to predict new outcomes and in this regard we consider the results of the confirmation study to be especially noteworthy. The correlation between ESP performance and extraversion in the confirmation study is very close to that predicted by the meta-analysis. This is the second test of the predictive validity of meta-analysis in parapsychological problem areas; we have previously reported that ESP ganzfeld performance in a new series of studies (Honorton, et al., 1990), closely matched the outcomes of earlier studies in a meta-analysis (Honorton, 1985). Predictability is the hallmark of successful science and these findings lead us to be optimistic concerning the prospect that parapsychology may be approaching this more advanced stage of development.

References

Briggs, K. C., and Myers, I. B. (1957). *Myers-Briggs Type Indicator Form F.* Palo Alto, CA: Consulting Psychologists Press, Inc.

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York: Academic Press. (Revised Edition.)

Eysenck, H. J. (1967). Personality and extra-sensory perception. *Journal of the Society for Psychical Research*, 44, 55-70.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for Meta-Analysis*. New York: Academic Press.

Hedges, L. V. (1987). How hard is hard science, how soft is soft science? The empirical cumulativeness of research. *American Psychologist*, 42, 443-455.

Honorton, C. (1985). Meta-analysis of psi ganzfeld research: a response to Hyman. *Journal of Parapsychology*, 49, 51-92.

Honorton, C., Berger, R. E., Varvoglis, M. P., Quant, M., Derr, P., Hansen, G., Schechter, E. I., and Ferrari, D. C. (1990). Psi communication in the ganzfeld: experiments with an automated testing system and a comparison with a meta-analysis of earlier studies. In *Research in Parapsychology 1989*. Metuchen, NJ: Scarecrow Press. (In press.)

Honorton, C., & Ferrari, D.C. (1989) "Future Telling": a meta-analysis of forced-choice precognition experiments, 1935-1987. *Journal of Parapsychology*, 53, in press.

Hyman, R. (1985). The psi ganzfeld experiment: A critical appraisal. *Journal of Parapsychology*, 49, 3-49.

McCarthy, D., & Schechter, E. I. (1986). Estimating effect size from critical ratios. In D. H. Weiner & D. I. Radin (Eds.) *Research in Parapsychology 1985*. Scarecrow Press, pp. 95-96.

Palmer, J. (1977). Attitudes and personality traits in experimental ESP research. In B. B. Wolman (Ed.) *Handbook of parapsychology*. New York: Van Nostrand Reinhold.

Palmer, J., & Lieberman, R. (1975). The influence of psychological set on ESP and out-of-the-body experiences. *Journal of the American Society for Psychical Research*, 69, 193-213.

Radin, D. I., & Nelson, R. D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics*, 19, 1499-1514.

Rosenthal, R. (1984). *Meta-Analytic procedures for social research*. Beverly Hills, CA: Sage.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Rosenthal, R., & Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences*, 3, 377-386.

Sargent, C. L. (1981). Extraversion and performance in 'extra-sensory perception' tasks. *Personality and Individual Differences*, 2, 137-143.

Stanford, R. G., and Sargent, C. L. (1983). Z scores in free-response methodology: comments on their utility and correction of an error. *Journal of the American Society for Psychical Research*, 77, 319-326.

Super, D. E. (1942). The Bernreuter Personality Inventory: a review of research. *Psychological Bulletin*, 39, 94-125.

Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Studies Used in the Meta-Analysis

Some reports contain more than one study. For reports with multiple studies, the number of studies is indicated in brackets following the reference.

Ashton, H. T., Dear, P. R., & Harley, T. A. (1981). A four-subject study of psi in the ganzfeld. *Journal of the Society for Psychical Research*, 51, 12-21.

Astrom, J. (1965). GESP and the MPI measures. *Journal of Parapsychology*, 29, 292-293.

Bellis, J., & Morris, R. L. (1980). Openness, closedness and psi. *Research in Parapsychology 1979*, 98-99.

Braud, L. W. (1976). Openness versus closedness and its relationship to psi. *Research in Parapsychology 1975*, 155-159.

Braud, L. W. (1977). Openness vs. closedness and its relationship to psi. *Research in Parapsychology 1976*, 162-165.

Casper. G. W. (1952). Effects of the receiver's attitude toward the sender in ESP tests. *Journal of Parapsychology*, 16, 212-218.

Fisk, G. W (1960). The Rhodes experiment. Linkage in extra-sensory perception by M. C. Marsh. *Journal of the Society for Psychical Research*, 40, 219-239 and M. C. Marsh. (unpublished). *Linkage in Extra-Sensory Perception*. Unpublished doctoral dissertation, Dept. of Psychology, Rhodes University, Grahamstown, South Africa. 450 pages.

Green, C. E. (1966). Extra-sensory perception and the Maudsley Personality Inventory. *Journal of the Society for Psychical Research*, 43, 285-286.

Green, C. E. (1966). Extra-sensory perception and the extraversion scale of the Maudsley Personality Inventory. *Journal of the Society for Psychical Research*, 43, 337.

Haraldsson, E. (1970). Psychological variables in a GESP test using plethysmograph recordings. *Proceedings of the Parapsychological Association*. 7, 6-7.

Harley, T. A., & Sargent, C. L. (1980). Trait and state factors influencing ESP performance in the ganzfeld. *Research in Parapsychology 1979*, 126-127.

Humphrey, B. M. (1945). An exploratory correlation study of personality measures and ESP scores. *Journal of Parapsychology*, 9, 116-123. [3 studies]

Humphrey, B. M. (1951). Introversion-extraversion ratings in relation to scores in ESP tests. *Journal of Parapsychology*, 15, 252-262.

Kanthamani, B. K. (1966). ESP and social stimulus. *Journal of Parapsychology*, 30, 31-38.

Kanthamani, B. K., & Rao, K. R. (1972). Personality characteristics of ESP subjects: III. Extraversion and ESP. *Journal of Parapsychology*, 36, 198-212.

Krishna, S. R., & Rao, K. R. (1981). Personality and 'belief' in relation to language ESP scores. *Research in Parapsychology 1980*, 61-63. [2 studies]

McElroy, W. A., and Brown, W. K. R. (1950 Electric shocks for errors in ESP card tests. *Journal of Parapsychology*, 14, 257-266.

Nash, C. B. (1966). Relation between ESP scoring level and the Minnesota Multiphasic Personality Inventory. *Journal of the American Society for Psychical Research*, 60, 56-62. [8 studies]

Nicol, J. F., & Humphrey, B. M. (1953). The exploration of ESP and human personality. *Journal of the American Society for Psychical Research*, 47, 133-178.

Nicol, J. F., & Humphrey, B. M. (1955). The repeatability problem in ESP-personality research. *Journal of the American Society for Psychical Research*, 49, 125-156.

Nielsen, W. (1970). Relationships between precognition scoring level and mood. *Journal of Parapsychology*, 34, 93-116.

Nielsen, W. (1970). Studies in group targets: a social psychology class. *Proceedings of the Parapsychological Association*, 7, 55-57.

Nielsen, W. (1970). Studies in group targets: an unusual high school group. *Proceedings of the Parapsychological Association*, 7, 57-58.

Sargent, C. L. (1978). Hypnosis as a psi-conducive state: a controlled replication study. *Journal of Parapsychology*, 42, 257-275. [2 studies]

Sargent, C. L. (1980). *Exploring psi in the ganzfeld*. New York: Parapsychology Foundation, Inc. [2 studies]

Sargent, C. L., Bartlett, H. J., and Moss, S. P. (1982). Response structure and temporal incline in ganzfeld free-response GESP testing. *Journal of Parapsychology*, 46, 85-110. [2 studies]

Sargent, C. L., Harley, T. A., Lane, J., & Radcliffe, K. (1981). Ganzfeld psi-optimization in relation to session duration. *Research in Parapsychology 1980*, 82-84.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Sargent, C. L., & Harley, T. A. (1981). Three studies using a psi-predictive trait variable questionnaire. *Journal of Parapsychology*, 45, 199-214.

Sargent, C. L., and Matthews, G. (1982). Ganzfeld GESP performance with variable-duration testing. *Research in Parapsychology 1981*, 159-160.

Shields, E. (1962). Comparison of children's guessing ability (ESP) with personality characteristics. *Journal of Parapsychology*, 26, 200-210. [2 studies]

Shrager, E. F. (1978). The effects of sender-receiver relationship and associated personality variables on ESP scores. *Journal of the American Society for Psychical Research*, 72, 35-47. [2 studies]

Szczygielski, D., & Schmeidler, G. R. (1975). ESP and two measures of introversion. *Research in Parapsychology 1974*, 15-17.

Thalbourne, M. A., Beloff, J., and Delanoy, D. (1982). A test for the 'extraverted sheep versus introverted goats' hypothesis. *Research in Parapsychology 1981*, 155-156. [2 studies]

Thalbourne, M. A., Beloff, J., Delanoy, D., & Jungkuntz, J. H. (1983). Some further tests of the extraverted sheep versus introverted goats hypothesis. *Research in Parapsychology 1982*, 199-200. [4 studies]

Thalbourne, M. A., and Jungkuntz, J. H. (1983). Extraverted sheep versus introverted goats: experiments VII and VIII. *Journal of Parapsychology*, 47, 49-51. [2 studies]

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Journal of Parapsychology, Vol. 50, December 1986

META-ANALYTIC PROCEDURES AND THE NATURE OF REPLICATION: THE GANZFELD DEBATE

BY ROBERT ROSENTHAL

ABSTRACT: This paper is a commentary on the valuable debate between Charles Honorton (1985) and Ray Hyman (1985) about the evidence for psi in the ganzfeld situation. Their debate was a creative, constructive, and task-oriented dialogue that served admirably to sharpen the issues involved. In my commentary I focus on the concept of replication, distinguishing the troublesome older view with a more useful alternative. Specific issues related to replication are discussed including problems of multiple testing, subdividing studies, weighting replications, and problems of small effects. The earlier meta-analytic work is summarized, evaluated, and compared with a meta-analysis of a different controversial area. Rival hypotheses of procedural and statistical types are discussed, and a tentative inference is offered. The conclusion calls for wider use of newer views of the success of replication.

Science in general and parapsychological inquiry in particular have been well served by the recent ganzfeld debate between Charles Honorton (1985) and Ray Hyman (1985) as organized by the *Journal*'s editor, K. Ramakrishna Rao. Two serious and highly knowledgeable scholars have invested a great amount of time, energy, and creative thought to produce a debate that is a model of task-oriented, constructive dialogue. It is clear that the participants have been devoted to clarifying and understanding the scientific issues rather than simply to "scoring points."

As a result of their efforts we have an excellent review of the issues to be considered in evaluating the data generated by the ganzfeld experiments. In addition, through their meta-analytic work, we have an enormously valuable quantitative summary of the ganzfeld studies. In the end, Hyman and Honorton have not resolved all their differences, nor is it likely that they will. Hyman has raised cogent and telling questions. Honorton has answered them in cogent and telling terms. I am sure that Hyman will have excellent

The preparation of this paper and the development of some of the procedures described within it were supported by the National Science Foundation. Much of the summary and interpretation of the meta-analyses will be included in a paper commissioned by the National Academy of Sciences that is in preparation by Monica J. Harris and Robert Rosenthal.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

study failed to replicate that of Smith. Such errors are made very frequently in most areas of psychology and the other behavioral sciences.

Pseudo-Successful Replications

Return now to Table 1 and focus attention on cell B, the cell of "successful replication." Suppose that two investigators both rejected the null hypothesis at $p < .05$ with both results in the same direction. Suppose further, however, that in one study the effect size r was .90 whereas in the other study the effect size r was only .10, significantly smaller than the r of .90 (Rosenthal & Rubin, 1982a). In this case our interpretation is more complex. We have indeed had a successful replication of the rejection of the null, but we have not come even close to a successful replication of the effect size.

"Successful Replication" of Type II Error

Cell C of Table 1 represents the situation in which both studies failed to reject the null hypothesis. Under those conditions investigators might conclude that there was no relationship between the variables investigated. Such a conclusion could be very much in error, the more so the lower the power of the two studies was low (Cohen, 1977). If power levels of the two studies (assuming medium effect sizes in the population) were very high, say .90 or .95, then two failures to obtain a significant relationship would provide evidence that the effect investigated was not likely to be a very large effect. If power calculations had been made assuming a very small effect size, two failures to reject the null although not providing strong evidence for the null would at least suggest that the size of the effect in the population was probably quite modest.

If sample sizes of the two studies failing to reject the null were modest so that power to detect all but the largest effects were low, very little could be concluded from two failures to reject except that the effect sizes were unlikely to be enormous. For example, two investigators with N's of 20 and 40, respectively, find results not significant at $p < .05$. The effect sizes *phi* (i.e., r for dichotomous variables) were .29 and .20, respectively, and both p's were approximately .20. The combined p of these two results, however, is $.035[(z_1 + z_2)/\sqrt{2} = z]$, and the mean effect size in the mid-.20's is not trivial (Rosenthal & Rubin, 1982b).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 3
COMPARISON OF TWO SETS OF REPLICATIONS

	Replication sets			
	A		B	
	Study 1	Study 2	Study 1	Study 2
N	96	15	98	27
p (two-tailed)	.05	.05	.01	.18
z (p)	1.96	1.96	2.58	1.34
r	.20	.50	.26	.26
$z(r)$	.20	.55	.27	.27
Cohen's q ($z_{r_1} - z_{r_2}$)	.35		.00	

Comparing Views of Replication

The traditional, not very useful, view of replication modeled in Table 1 has two primary characteristics:

1. It focuses on significance level as the relevant summary statistic of a study.
2. It makes its evaluation of whether replication has been successful in a dichotomous fashion. For example, replications are successful if both or neither $p < .05$ (or .01, etc.), and they are unsuccessful if one $p < .05$ (or .01, etc.) and the other $p > .05$ (or .01, etc.). Psychologists' reliance on a dichotomous decision procedure accompanied by an untenable discontinuity of credibility in results varying in p levels has been well documented (Nelson, Rosenthal, & Rosnow, 1986; Rosenthal & Gaito, 1963, 1964).

The newer, more useful view of replication success has two primary characteristics:

1. It focuses on effect size as the more important summary statistic of a study with only a relatively minor interest in the statistical significance level.
2. It makes its evaluation of whether replication has been successful in a continuous fashion. For example, two studies are not said to be successful or unsuccessful replicates of each other but, rather, the degree of failure to replicate is specified.

Table 3 shows two sets of replications. Replication set A shows two results both rejecting the null but with a difference in effect sizes of .30 in units of r or .35 in units of Fisher's z transformation of r (Cohen, 1977; Rosenthal & Rosnow, 1984; Snedecor & Cochran, 1980). That difference, in units of r or Fisher's z is the degree

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

multiple questions, multiple dependent variables make good scientific sense. However, as both Honorton (1985) and Hyman (1985) point out, the use of multiple dependent variables may affect the accuracy of the p levels computed. For example, if five dependent variables are used and one of these is found to show an effect at p = .05, it would be misleading to say that an effect has been demonstrated at $p < .05$. That is because the actual p of finding *one* p significant at .05 (or any other chosen level) increases as the number of tests made increases. That is not a good reason to decrease the variety of dependent variables used, assuming there is a good theoretical basis for choosing to use each one.

Alternate procedures are available. Bonferroni procedures can be used to adjust for the number of tests made (Rosenthal & Rubin, 1983). To overcome the conservatism of this basic approach and decrease Type II errors, it is possible to weight the dependent variables according to their importance and apply a so-called ordered Bonferroni procedure (Rosenthal & Rubin, 1984, 1985). Perhaps it is most useful, however, to apply specially developed procedures that integrate all the information from all the dependent variables and obtain only a single overall test of significance and effect size estimate. This can be accomplished very easily so long as we have reasonable estimates of the intercorrelations among the dependent variables (Rosenthal & Rubin, 1986).

Subdividing Studies

An issue discussed in the ganzfeld debate has to do with the subdivision of studies into substudies as a function of different experimental procedures or individual difference variables such as sex, age, degree of belief in psi effects, and the like (Schmeidler, 1968). As long as all the data are preserved and entered into the meta-analysis, no harm is done by subdividing. Indeed, subdividing is very useful in the search for moderator variables (Rosenthal, 1984).

Subdividing could have a very biasing effect on the accuracy of a cited p value if the overall data are subdivided in various ways, significant results are reported for one or more substudies, and the rest of the substudies are "thrown away." In the ordinary more proper application of meta-analytic procedures, however, subdividing makes little difference. Consider a psi experiment with an overall nonsignificant effect (p = .13, two-tailed). After the study is over, it is noted that about half the subjects were favorable toward psi and half were not and that there had been both female and male sub-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 4
SUBDIVISION OF A LARGER EXPERIMENT

	Believing subjects		Disbelieving subjects	
	Two-tailed p	z	Two-tailed p	z
Females	.05	2.0	.62	0.5
Males	.32	1.0	.62	−0.5

Note: For the study as a whole, p was .13 and z was 1.5 before subdividing. Positive z's reflect results in the predicted direction; negative z's reflect results in the unpredicted direction.

jects. Suppose that a subgroup of subjects, say female believers, show a significant psi effect but the remaining groups do not. No harm is done by reporting that fact, though an adjustment is useful in reporting the obtained p that takes into account how many subgroups were tested. It is essential, however, that the results of significance tests for the nonsignificant subgroups also be entered into the meta-analysis.

Table 4 illustrates the situation; four substudies have been formed, only one of which was significant. When we combine the results of the four substudies, however, we find the overall z to be $[(2.0) + (1.0) + (0.5) + (-0.5)]/\sqrt{4} = 1.5, p = .13$, two-tailed. Essentially, subdividing makes little difference so long as no data are discarded. If a particular substudy showed great promise of evidencing psi, nothing would prevent the investigator from conducting new studies using only the preselected experimental conditions or types of subjects. It would also be appropriate to conduct a meta-analysis on all the substudies that could be found that met the promising condition. In that case, however, the initial "study of discovery" should be entered with an adjustment for the fact that several tests of significance were computed (Rosenthal & Rubin, 1983, 1984).

Flaw Effects and Weighting Replications

There are few flawless studies in the behavioral sciences. Flaws can increase Type I or Type II errors, and the wise meta-analyst would do well to note how well Hyman (1985) and Honorton (1985) have searched for and evaluated flaws. For each flaw, it would be desirable to make some estimate of how much difference it made to the outcome. In the present debate some flaws seemed to make a difference and others did not. When flaws matter we can adjust for

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

of failure to replicate. That both studies were able to reject the null and at exactly the same p level is simply a function of sample size. Replication set B shows two studies with different p values, one significant at $< .05$, the other not significant. However, the two effect size estimates are in excellent agreement. We would say, accordingly, that replication set B shows more successful replication than does replication set A.

It should be noted that the values of Table 3 were chosen so that the combined probability of the two studies of set A would be identical to the combined probability of the two studies of set B; $(z_1 + z_2)/\sqrt{2} = z$ of 2.77, $p = .0028$, one-tailed.

The Metrics of the Success of Replication

Once we adopt a view of the success of replication as a function of similarity of effect sizes obtained, we can become more precise in our assessments of the success of replication. Figure 1 shows the "replication plane" generated by crossing the results of the first study conducted (expressed in units of the effect size r) by the results of the second study conducted. All perfect replications, those in which the effect sizes are identical in the two studies, fall on a diagonal rising from the lower left corner $(-1.00, -1.00)$ to the upper right corner $(+1.00, +1.00)$. The results of replication set B from Table 3 are shown to fall exactly on the diagonal of successful replication $(+.26, +.26)$. The results of replication set A are shown to fall somewhat above the line representing perfect replication. Figure 1 shows that although set B reflects a more successful replication than set A, the latter is also located fairly close to the line and is, therefore, a fairly successful replication set as well.

Cohen's q. An alternative to the indexing of the success of replication by the difference between obtained effect size r's is to transform the r's to Fisher's z's before taking the difference. Fisher's z metric is distributed nearly normally and can thus be used in setting confidence intervals and testing hypotheses about r's, whereas r's distribution is skewed, and the more so as the population value of r moves further from zero. Cohen's q is especially useful for testing the significance of difference between two obtained effect size r's. This is accomplished by means of the fact that

$$q \bigg/ \sqrt{\frac{1}{N_1 - 3} + \frac{1}{N_2 - 3}}$$

is distributed as z, the standard normal deviate (Rosenthal, 1984;

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

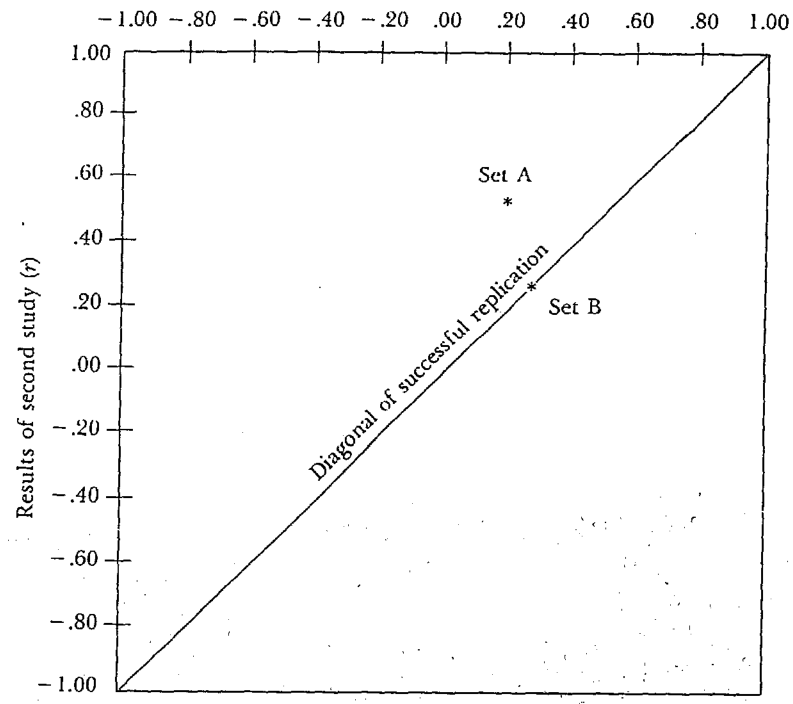


Figure 1. The replication plane.

Rosenthal & Rubin, 1982a; Snedecor & Cochran, 1980). When there are more than two effect size r's to be evaluated for their variability (i.e., heterogeneity), the three references above all provide the appropriate formula for computing the test of the heterogeneity of r's.

Issues Related to Replication

Multiple Testing

In ganzfeld studies, in parapsychological research more broadly, and, indeed, in most areas of behavioral science, it is common that more than one test of significance is computed to evaluate a research hypothesis. There may, for example, be a set of several dependent variables used to evaluate outcome. So long as there are

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

these flaws in our weighting of studies. For example, we can give weights of zero to truly terrible studies and lowered but nonzero weights to less than truly terrible studies. Such weighting may lead to less biased conclusions than simple discarding of studies for flaws (Fiske, 1978; Rosenthal, 1984; Rosenthal & Rubin, 1985).

Replication Difficulty and Small Effects

Although Hyman (1985) and Honorton (1985) disagree on the degree of confidence warranted by the ganzfeld literature, they agree that the results reported do not reflect an enormous magnitude of effect. In Cohen's (1977) terminology, the average size of the ganzfeld effect reported by Hyman (1985) and Honorton (1985) is on the small side. That, of course, is not surprising. Controversial research areas are characterized by small effect sizes. For example, in a recent review of five controversial areas of human performance research, Harris and Rosenthal (1986) estimated the actual effect sizes (r) to range only from .00 to .18 with a median of .10 and a 95% confidence interval ranging from .02 to .19.

Small effect sizes are just what we should expect from controversial areas. According to fundamental principles of statistical power (Cohen, 1977), if the true effect size were substantial, studies with only modest sample sizes would routinely be able to reject the null. For example, if the population value of r were .60, 90% of replication attempts would be significant at $p < .05$ with sample sizes of 24 (Cohen, 1977, p. 92). However, if the population value of r were .10, the median of our five controversial areas (Harris & Rosenthal, 1986), only 7% of replication attempts would be significant at $p < .05$ with sample sizes of 24. For the small population value of r (.10), it would require sample sizes of over 1,000 to achieve a 90% rate of rejecting the null at $p < .05$.

Even though controversial research areas are characterized by small effects (including zero as a possibility), that does not mean that the effects are of no practical importance. Indeed, the median small effect of five areas cited above ($r = .10$) is equivalent to improving our success rate from 45% to a success rate of 55% (Rosenthal & Rubin, 1982b).

Before leaving the topic of replication difficulty, it may help us to place this problem in useful perspective by noting that it is not only in the parapsychological or other behavioral sciences that replication difficulties emerge. Indeed, students of the physical sciences have pointed out failures to replicate the construction of TEA-lasers

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

despite the availability of detailed instructions for replication. Apparently TEA-lasers could be replicated dependably only when the replication instructions were accompanied by a scientist who had actually built a laser (Collins, 1985).

Summarizing the Meta-Analyses

Hyman (1985) and Honorton (1985) have done important meta-analytic work on the topic of the ganzfeld experiments; it is this work I summarize here.

Five indices of "psi" success have been used in ganzfeld research (Honorton, 1985). One criticism of research in this area is that some investigators used several such indices in their studies and failed to adjust their reported levels of significance (p) for the fact that they had made multiple tests (Hyman, 1985). Because most studies used a particular one of these five methods, the method of direct hits, Honorton focused his meta-analysis on just those 28 studies (of a total of 42) for which direct hit data were available.

The method of direct hits scores a success only when the single correct target is chosen out of a set of t total targets. Thus, the probability of success on a single trial is $1/t$ with t usually $= 4$ but sometimes 5 or 6. The other methods, using some form of partial credit, appear to be more precise in that they use more of the information available. Although they differ in their interpretation of the results, Honorton (1985) and Hyman (1985) agree quite well on the basic quantitative results of the meta-analysis of these 28 studies. This agreement holds both for the estimation of statistical significance (Honorton, 1985, p. 58) and of effect size (Hyman, 1985, p. 13).

Stem-and-Leaf Display

Table 5 shows a stem-and-leaf display of the 28 effect size estimates based on the direct hits studies summarized by Honorton (1985, p. 84). The effect size estimates shown in Table 5 are in units of Cohen's h, which is the difference between (a) the arcsine transformed proportion of direct hits obtained and (b) the arcsine transformed proportion of direct hits expected under the null hypothesis (i.e., $1/t$). The advantage of h over j, the difference between raw proportions, is that all h values that are identical are identically detectable whereas all j values that are identical (e.g., $.65 - .45$ and $.25 - .05$) are not equally detectable (Cohen, 1977, p. 181).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 5
STEM-AND-LEAF PLOT OF "DIRECT HIT" GANZFELD STUDIES: COHEN'S *h*

Stem	Leaf
1.4	4
1.3	3
1.2	
1.1	
1.0	
.9	
.8	
.7	3
.6	
.5	8
.4	0 2 2 2 4
.3	1 2 2 4 4 7 8
.2	2
.1	3 8 8
.0	7 7 9
−.0	5
−.1	0
−.2	
−.3	2
−.4	0
−.5	
−.6	
−.7	
−.8	
−.9	3

Tukey (1977) developed the stem-and-leaf plot as a special form of frequency distribution to facilitate the inspection of a batch of data. Each number in the data batch is made up of one stem and one leaf, but each stem may serve several leaves. Thus, the stem .1 is followed by leaves of 3, 8, 8 representing the numbers .13, .18, .18. The first digit is the stem; the next digit is the leaf. The stem-and-leaf display functions as any other frequency distribution but the original data are retained precisely.

Distribution of studies. From Table 5 we see that the distribution of effect sizes is unimodal, with the bulk of the results (80%) falling between −.10 and .58. The distribution is nicely symmetrical, with the skewness index (g_1 = .17) only 24% of that required for significance at $p < .05$ (Snedecor & Cochran, 1980, pp. 78–79, 492). The tails of the distribution, however, are too long for normality with kurtosis index g_2 = 2.04, p = .02. Relative to what we would expect from a normal distribution, we have studies that show larger positive and larger negative effect sizes than would be reasonable. Indeed, the two largest positive effect sizes are significant outliers at $p < .05$, and the largest negative effect size approaches significance, with a Dixon index of .37 compared to one of .40 for the largest positive effect size (Snedecor & Cochran, 1980, pp. 279–280, 490). The total sample of studies is still small; however, if a much larger sample showed the same result, that would be a pattern consistent with the idea that both strong positive results ("psi") and strong negative results ("psi-missing") might be more likely to find their way into print or at least to be more available to a meta-analyst.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Distribution of subjects. It is useful to examine the distribution of effect sizes obtained in the summarized studies. It would also be useful to examine the distribution of effect sizes obtained by individual subjects *within* the studies summarized. For example, in a study with a mean *h* of .20, is the distribution of *h* fairly normal with centering at .20, or is the distribution skewed with the bulk of the subjects centered closer to zero but with a few subjects earning consistently high values of *h*?

Distribution of investigators. Just as it is useful to examine the distribution of the results of studies and of subjects within studies, it is also useful to examine the distribution of results obtained by different investigators (Honorton, 1985; Hyman, 1985; Rosenthal, 1969, 1984). The 28 direct hit studies were conducted by 10 different investigators (Honorton, 1985, p. 60). Four investigators conducted only one study each, two conducted two studies each, two conducted three studies each, one conducted five studies, and one conducted nine studies. Analysis of variance showed that these 10 investigators differed significantly and importantly in the average magnitude of the effects they obtained with $F(9,18)$ = 3.81, $p < .01$, *eta* = .81. Interestingly, there was little relationship between the mean effect size obtained by each investigator and the number of studies conducted (r = .11; $t(8)$ = 0.31, $p > .70$).

That different investigators may obtain significantly different results from their subjects is well known in various areas of psychology (Rosenthal, 1966). For example, in such a standard experimental area as eyelid conditioning, studies conducted at Iowa obtained results in the predicted direction 94% of the time, whereas those conducted elsewhere obtained such results only 62% of the time with $\chi^2(1)$ = 4.05, $p < .05$, N = 25, r = .40 (Rosenthal, 1966, p. 24; Spence, 1964).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 6
STATISTICAL SUMMARY OF "DIRECT HIT" GANZFELD STUDIES

Central tendency (Cohen's h)		*Variability*	
Unweighted mean	.28	Maximum	1.44
Weighted mean	.23	Quartile 3 (Q3)	.42
Median	.32	Median (Q2)	.32
Proportion positive sign	.82	Quartile 1 (Q1)	.08
		Minimum	−.93
Significance tests		Q3 − Q1	.34
Combined Stouffer z	6.60	$\hat{\sigma}$: [.75 (Q3 − Q1)]	.26
t test of mean z	3.23	S	.45
z of proportion positive	3.40		
		Correlation of h	
		With z	.86
		With raw j	.98

Confidence intervals[a]

	From	*To*
80%	.17	.39
95%	.11	.45
99%	.04	.52
99.9%	−.03	.59

[a]Based on *N* of 28 studies.

Summary of Stem-and-Leaf Display

Table 6 provides a summary of the stem-and-leaf display of Table 5 and some additional useful information about central tendency, variability, significance tests, confidence intervals, and correlations between Cohen's h and (a) significance level (z) and (b) raw difference in proportions (j). Only a few comments are required.

Effect size. The bulk of the results (82%) show a positive effect size where 50% would be expected under the null (p = .0004). The mean effect size, h, of .28 is equivalent to having a direct hit rate of .38 when .25 was expected under the null. The 95% confidence interval suggests the likely range of effect sizes to be from .11 to .45, equivalent to accuracy rates of .30 to .46 when .25 was expected under the null hypothesis.

Significance testing. The overall probability that obtained accuracy was better than the accuracy expected under the null was a p of $3.37/10^{11}$ associated with a Stouffer z of 6.60 (Mosteller & Bush, 1954; Rosenthal, 1978a, 1984).

File-drawer analysis. A combined p as low as that obtained can be used as a guide to the tolerance level for null results that never found their way into the meta-analytic data base (Rosenthal, 1979, 1984). It has long been believed that studies failing to reach statistical significance may be less likely to be published (Rosenthal, 1966; Sterling, 1959). Thus it may be that there is a residual of nonsignificant studies languishing in the investigators' file drawers. With simple calculations, it can be shown that, for the current studies summarized, there would have to be 423 studies with mean p = .50, one-tailed, or z = 0.00 in those file drawers before the overall combined p would become just > .05, as Honorton (1985) has pointed out.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

That many studies unretrieved seems unlikely for this specialized area of parapsychology (Honorton, 1985; Hyman, 1985). Based on experience with meta-analyses in other domains of research (e.g., interpersonal expectancy effects) the mean z or effect size for nonsignificant studies is not 0.00 but a value pulled strongly from 0.00 toward the mean z or mean effect size of the obtained studies (Rosenthal & Rubin, 1978).

Comparison with an Earlier Meta-Analysis

It is instructive to compare the results of the ganzfeld research meta-analysis by Honorton (1985) with the results of an older and larger meta-analysis of another controversial research domain—that of interpersonal expectancy effects (Rosenthal & Rubin, 1978). In that analysis, eight areas of expectancy effects were summarized; effect sizes (Cohen's d, roughly equivalent to Cohen's h) ranged from .14 to 1.73 with a grand mean d of .70. Honorton's mean effect size (h = .28) exceeds the mean d of two of the eight areas (reaction time experiments [d = .17], and studies using laboratory interviews [d = .14]).

The earlier meta-analysis displayed the distribution of the z's associated with the obtained p levels. Table 7 shows a comparison of the two meta-analyses' distributions of z's. It is interesting to note the high degree of similarity in the distributions of significance levels. The total proportion of significant results is somewhat higher for the ganzfeld studies but not significantly so ($\chi^2(1)$ = 1.07, N = 373, p = .30, ϕ = .05).

INTERPRETING THE META-ANALYTIC RESULTS

Although the results of the meta-analysis are clear, the meaning of these results is open to various interpretations. The most obvious

TABLE 7
PROPORTION OF STUDIES REACHING CRITICAL LEVELS OF SIGNIFICANCE FOR TWO RESEARCH AREAS

Interval for z	Expected proportion	Expectancy research[a]	Ganzfeld research[b]	Difference
Predicted direction				
+ 3.72 and above	.0001	.07	.04	−.03
+ 3.09 and above	.001	.12	.18	.06
+ 2.33 and above	.01	.19	.25	.06
+ 1.65 and above	.05	.36	.43	.07
Not significant				
− 1.64 to + 1.64	.90	.60	.50	−.10
Unpredicted direction				
− 1.65 and below	.05	.03	.07	.04

[a]N = 345 studies; from Rosenthal & Rubin (1978).
[b]N = 28 studies; from Honorton (1985).

interpretation might be that at a very low p, and with a fairly impressive effect size, the ganzfeld psi phenomenon has been demonstrated. However, there are rival hypotheses that will need to be considered, many of them put forward in the detailed evaluation by Hyman (1985).

Procedural Rival Hypotheses

Sensory leakage. A standard rival hypothesis to the hypothesis of ESP is that sensory leakage occurred and that the receiver was knowingly or unknowingly cued by the sender or by an intermediary between the sender and receiver. As early as 1895, Hansen and Lehmann (1895) described "unconscious whispering" in the laboratory, and Kennedy (1938, 1939) was able to show that senders in telepathy experiments could give auditory cues to their receivers quite unwittingly. Ingenious use of parabolic sound reflectors made this demonstration possible. Moll (1898), Stratton (1921), and Warner and Raible (1937) all gave early warnings on the dangers of unintentional cueing (for summaries see Rosenthal, 1965, 1966). The subtle kinds of cues described by these early workers were just the kind we have come to look for in searching for cues given off by experimenters that might serve to mediate the experimenter expectancy effects found in laboratory settings (Rosenthal, 1966, 1985).

By their nature, ganzfeld studies tend to minimize problems of sensory cueing. An exception occurs when the subject is asked to choose which of four (or more) stimuli has been "sent" by another person or agent. When the same stimuli held originally by the sender are shown to the receiver, finger smudges or other marks may serve as cues. Honorton has shown, however, that studies controlling for this type of cue yield at least as many significant effects as do the studies not controlling for this type of cue.

Recording errors. A second rival hypothesis has nearly as long a history. Kennedy and Uphoff (1939) and Sheffield and Kaufman (1952) both found biased errors of recording the data of parapsychological experiments. In a meta-analysis of 139,000 recorded observations in 21 studies, it was found that about 1% of all observations were in error and that, of the errors committed, twice as many favored the hypothesis as opposed it (Rosenthal, 1978b). Although it is difficult to rule recording errors out of ganzfeld studies (or any other kind of research), their magnitude is such that they could probably have only a small biasing effect on the estimated average effect size (Rosenthal, 1978b, p. 1007).

Intentional error. The very recent history of science has reminded us that even though fraud in science is not quite of epidemic proportion, it must be given close attention (Broad & Wade, 1982; Zuckerman, 1977). Fraud in parapsychological research has been a constant concern, a concern found to be justified by periodic flagrant examples (Rhine, 1975). In the analyses of Hyman (1985) and Honorton (1985), in any case, there appeared to be no relationship between degree of monitoring of participants and the results of the study.

Statistical Rival Hypotheses

File-drawer issues. The problem of biased retrieval of studies for any meta-analysis was described earlier. Part of this problem is addressed by the 10-year-old norm of the Parapsychological Association of reporting negative results at its meetings and in its journals (Honorton, 1985). Part of this problem is addressed also by Blackmore (1980), who conducted a survey to retrieve unreported ganzfeld studies. She found that 7 of her total of 19 studies were judged significant overall by the investigators. This proportion of significant results (.37) was not significantly (or appreciably) lower than the proportion of published studies found significant (.43) in Honorton's (1985) meta-analysis of direct hit ganzfeld studies ($\chi^2(1)$ =

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4
Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

0.17, ϕ = .06. Somewhat similar results were obtained by Sommer (in press) in her analysis of research on the menstrual cycle. She found 61% of the published results to be significant compared to 40% of the unpublished studies; $\chi^2(1) = 2.30$, $p < .065$, one-tailed, $\phi = .20$. The results of the Blackmore and Sommer studies did not differ significantly ($z = 0.69$). Taken together, these studies provide only modest evidence for a serious file-drawer problem.

A problem that seems to be a special case of the file-drawer problem was pointed out by Hyman (1985). That was a possible tendency to report the results of pilot studies along with subsequent significant results when the pilot data were significant. At the same time it is possible that pilot studies were conducted without promising results, pilot studies that then found their way into the file drawers. In any case, it is nearly impossible to have an accurate estimate of the number of unretrieved studies or pilot studies actually conducted. Chances seem good, however, that there would be fewer than the 423 results of mean $z = 0.00$ required to bring the overall combined p to > .05.

Multiple testing. Each ganzfeld study may have more than one dependent variable for scoring degree of success. If investigators use these dependent variables sequentially until they find one significant at $p < .05$, the true p will be higher than .05 (Hyman, 1985). This issue was discussed earlier; it is not an inherently intractable one (Rosenthal & Rubin, 1986).

Randomization. Hyman (1985) has noted that the target stimulus may not have been selected in a truly random way from the pool of potential targets. To the extent that this is the case, the p values calculated can be in error. Hyman (1985) and Honorton (1985) disagree over the frequency in this sample of studies of improper randomization. In addition, they disagree over the magnitude of the relationship between inadequate randomization and study outcome. Hyman felt this relationship to be significant and positive; Honorton felt this relationship to be nonsignificant and negative. Because the median p level of just those 16 studies using random number tables or generators ($z = .94$) was essentially identical to that found for all 28 studies, it seems unlikely that poor randomization procedures were associated with much of an increase in significance level (Honorton, 1985, p. 71).

Statistical errors. Hyman (1985) and Honorton agree that 6 of the 28 studies contained statistical errors. However, the median effect size of these studies ($h = .33$) was very similar to the overall median ($h = .32$), so that it seems unlikely that these errors had a major

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

effect on the overall effect size estimate. Omitting these six studies from the analysis decreases the mean h from .28 to .26. Such a drop is equivalent to a drop of the mean accuracy rate from .38 to .37 when .25 is the expected value under the null.

A Tentative Inference

On the basis of the preceding summary and the very valuable meta-analytic evaluations of Honorton (1985) and Hyman (1985), what are we to believe? It would be easiest to say, "Let's wait until more data have been accumulated from studies purged of the problems noted by Hyman, Honorton, and others." That is not a realistic approach. At any point in time some judgment can be made, and though our judgment might be more accurate later on when those more nearly perfect studies become available, the situation for the ganzfeld domain seems reasonably clear. We feel it would be implausible to entertain the null given the combined p from these 28 studies. Given the various problems or flaws pointed out by Hyman and Honorton, the true effect size is almost surely smaller than the mean h of .28 equivalent to a mean accuracy of 38% when 25% is expected under the null. We are persuaded that the net result of statistical errors was a biased increase in estimated effect size of at least a full percentage point (from 37% to 38%). Furthermore, we are persuaded that file-drawer and related problems are such that some of the smaller effect size results have probably been kept off the market. If pressed to estimate a more accurate effect size, we might think in terms of a shrinkage of h from the obtained value of .28 to perhaps an h of .18. Thus, when the accuracy rate expected under the null is 1/4, we might estimate the obtained accuracy rate to be about 1/3.

Conclusion

Parapsychologists in particular and scientists in general owe a great debt of gratitude to Ray Hyman (1985) and Charles Honorton (1985) for their careful and extensive analytic and meta-analytic work on the ganzfeld problem. Their debate has yielded an especially high light/heat ratio, and many of the important issues have now been brought out into bold relief.

In my commentary on the ganzfeld debate, I focused most closely on the concept of replication. That seemed appropriate, not

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

only because of the centrality of the problem of replicability in the parapsychological literature, but also because of the centrality of the problem in many sciences, especially when the effect sizes sought in the population are small. The effect size zero is only a special case of the class of small effect sizes.

In closing I want only to suggest that parapsychological and other behavioral sciences would be well served to modify their view of the success of replication in the direction of the following newer view:

1. A replication is successful to the degree that the second study obtains an effect size similar to the effect size of the first study.

2. Three or more investigations are successful replicates of one another to the extent that the effect sizes are homogeneous.

3. Significance testing has nothing to do with success of replication though it can be useful in many ways, including the assessment of the likelihood of the null given all prior research (weighted as desired and as reasonable) and the likelihood of real differences among the effect sizes of two or more studies.

References

BLACKMORE, S. (1980). The extent of selective reporting of ESP ganzfeld studies. *European Journal of Parapsychology*, **3**, 213–219.

BROAD, W., & WADE, N. (1982). *Betrayers of the truth*. New York: Simon and Schuster.

COHEN, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

COLLINS, H. M. (1985). *Changing order: Replication and induction in scientific practice*. Beverly Hills, CA: Sage.

FISKE, D. W. (1978). The several kinds of generalization. *The Behavioral and Brain Sciences*, **3**, 393–394.

HANSEN, F. C. C., & LEHMANN, A. (1895). Ueber Unwillkürliches Flüstern. *Philosophische Studien*, **11**, 471–530.

HARRIS, M. J., & ROSENTHAL, R. (1986). *Interpersonal expectancy effects and human performance research*. Report prepared for the National Academy of Sciences.

HONORTON, C. (1985). Meta-analysis of psi ganzfeld research: A response to Hyman. *Journal of Parapsychology*, **49**, 51–91.

HYMAN, R. (1985). The ganzfeld psi experiment: A critical appraisal. *Journal of Parapsychology*, **49**, 3–49.

KENNEDY, J. L. (1938). Experiments on "unconscious whispering." *Psychological Bulletin*, **35**, 526. (Abstract)

KENNEDY, J. L. (1939). A methodological review of extra-sensory perception. [illegible]logic[illegible]lletin [illegible]59–1[illegible]

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

KENNEDY, J. L., & UPHOFF, H. F. (1939). Experiments on the nature of extra-sensory perception: III. The recording error criticism of extra-chance scores. *Journal of Parapsychology*, **3**, 226–245.

MOLL, A. (1898). *Hypnotism* (4th ed.). New York: Scribner.

MOSTELLER, F. M., & BUSH, R. R. (1954). Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology: Vol. 1. Theory and method* (pp. 289–334). Cambridge, MA: Addison-Wesley.

NELSON, N., ROSENTHAL, R., & ROSNOW, R. L. (1986). Interpretation of significance levels and effect sizes by psychological researchers. *American Psychologist*, **41**, 1299–1301.

RAO, K. R. (1985). The ganzfeld debate. *Journal of Parapsychology*, **49**, 1–2.

RHINE, J. B. (1975). Second report on a case of experimenter fraud. *Journal of Parapsychology*, **39**, 306–325.

ROSENTHAL, R. (1965). Clever Hans: A case study of scientific method. In O. Pfungst, *Clever Hans* (pp. ix–xlii). New York: Holt, Rinehart and Winston.

ROSENTHAL, R. (1966). *Experimenter effects in behavioral research*. New York: Appleton-Century-Crofts.

ROSENTHAL, R. (1969). Interpersonal expectations. In R. Rosenthal & R. L. Rosnow (Eds.), *Artifact in behavioral research* (pp. 181–277). New York: Academic Press.

ROSENTHAL, R. (1978a). Combining results of independent studies. *Psychological Bulletin*, **85**, 185–193.

ROSENTHAL, R. (1978b). How often are our numbers wrong? *American Psychologist*, **33**, 1005–1008.

ROSENTHAL, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, **86**, 638–641.

ROSENTHAL, R. (1984). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.

ROSENTHAL, R. (1985). Nonverbal cues in the mediation of interpersonal expectancy effects. In A. W. Siegman & S. Feldstein (Eds.), *Multichannel integrations of nonverbal behavior* (pp. 105–128). Hillsdale, NJ: Lawrence Erlbaum Associates.

ROSENTHAL, R., & GAITO, J. (1963). The interpretation of levels of significance by psychological researchers. *Journal of Psychology*, **55**, 33–38.

ROSENTHAL, R., & GAITO, J. (1964). Further evidence for the cliff effect in the interpretation of levels of significance. *Psychological Reports*, **15**, 570.

ROSENTHAL, R., & ROSNOW, R. L. (1984). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

ROSENTHAL, R., & RUBIN, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *The Behavioral and Brain Sciences*, **3**, 377–386.

ROSENTHAL, R., & RUBIN, D. B. (1979). Comparing significance levels of independent studies. *Psychological Bulletin*, **86**, 1165–1168.

ROSENTHAL, R., & RUBIN, D. B. (1982a). Comparing effect sizes of independent studies. *Psychological Bulletin*, **92**, 500–504.

ROSENTHAL, R., & RUBIN, D. B. (1982b). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, **74**, 166–169.

ROSENTHAL, R., & RUBIN, D. B. (1983). Ensemble-adjusted *p* values. *Psychological Bulletin*, **94**, 540–541.

ROSENTHAL, R., & RUBIN, D. B. (1984). Multiple contrasts and ordered Bonferroni procedures. *Journal of Educational Psychology*, **76**, 1028–1034.

ROSENTHAL, R., & RUBIN, D. B. (1985). Statistical analysis: Summarizing evidence versus establishing facts. *Psychological Bulletin*, **97**, 527–529.

ROSENTHAL, R., & RUBIN, D. B. (1986). Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin*, **99**, 400–406.

SCHMEIDLER, G. R. (1968). Parapsychology. In *International Encyclopedia o the Social Sciences* (pp. 386–399). New York: MacMillan & Free Press.

SHEFFIELD, F. D., KAUFMAN, R. S., & RHINE, J. B. (1952). A PK experimen at Yale starts a controversy. *Journal of the American Society for Psychica Research*, **46**, 111–117.

SNEDECOR, G. W., & COCHRAN, W. G. (1980). *Statistical methods* (7th ed.) Ames: Iowa State University Press.

SOMMER, B. (in press). The file drawer effect and publication rates in men strual cycle research. *Psychology of Women Quarterly*.

SPENCE, K. W. (1964). Anxiety (drive) level and performance in eyelid con ditioning. *Psychological Bulletin*, **61**, 129–139.

STERLING, T. D. (1959). Publication decisions and their possible effects o inferences drawn from tests of significance—or vice versa. *Journal c the American Statistical Association*, **54**, 30–34.

STRATTON, G. M. (1921). The control of another person by obscure sign *Psychological Review*, **28**, 301–314.

TRUZZI, M. (1981). Reflections on paranormal communication: A zetetic perspective. In T. A. Sebeok & R. Rosenthal (Eds.), *The Clever Ha phenomenon* (pp. 297–309). New York: New York Academy of Science

TUKEY, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-We ley.

WARNER, L., & RAIBLE, M. (1937). Telepathy in the psychophysical labor tory. *Journal of Parapsychology*, **1**, 44–51.

ZUCKERMAN, H. (1977). Deviant behavior and social control in science. I E. Sagarin (Ed.), *Deviance and social change* (pp. 87–138). Beverly Hil CA: Sage.

Department of Psychology
Harvard University
Cambridge, MA 02138

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

REPLICATION AND META-ANALYSIS IN PARAPSYCHOLOGY

Jessica Utts

Division of Statistics

University of California, Davis

1. INTRODUCTION

In a June 1990 Gallup Poll, 49% of the 1,236 respondents claimed to believe in extrasensory perception (ESP), and one in four claimed to have had a personal experience involving telepathy (Gallup and Newport, 1991). Other surveys have shown even higher percentages; the University of Chicago's National Opinion Research Council recently surveyed 1,473 adults, of which 67% claimed that they had experienced ESP (Greeley, 1987).

Public opinion is a poor arbiter of science, however, and experience is a poor substitute for the scientific method. For more than a century, small numbers of scientists have been conducting laboratory experiments to study phenomena such as telepathy, clairvoyance, and precognition, collectively known as "psi" abilities. This paper will examine some of that work, as well as some of the statistical controversies it has generated.

Parapsychology, as this field is called, has been a source of controversy throughout its history. Strong beliefs tend to be resistant to change even in the face of data, and many people, scientists included, seem to have made up their minds on the question without examining any empirical data at all. A critic of parapsychology recently acknowledged that "The level of the debate during the past 130 years has been an embarrassment for anyone who would like to believe that scholars and scientists adhere to standards of rationality and fair play" (Hyman, 1985a, p.89). While much of the controversy has focused on poor experimental design and

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

potential fraud, there have been attacks and defenses of the statistical methods as well, sometimes calling into question the very foundations of probability and statistical inference.

Most of the criticisms have been leveled by psychologists. For example, a 1988 report of the U.S. National Academy of Sciences concluded that "The committee finds no scientific justification from research conducted over a period of 130 years for the existence of parapsychological phenomena" (Druckman and Swets, 1988, p. 22). The chapter on parapsychology was written by a subcommittee chaired by a psychologist who had published a similar conclusion prior to his appointment to the committee (Hyman, 1985a, p.7). There were no parapsychologists involved with the writing of the report. Resulting accusations of bias (Palmer, Honorton and Utts, 1989) led U.S. Senator Claiborne Pell to request that the Congressional Office of Technology Assessment (OTA) conduct an investigation with a more balanced group. A one-day workshop was held on September 30, 1988 bringing together parapsychologists, critics, and experts in some related fields (including the author of this paper). The report concluded that parapsychology needs "a fairer hearing across a broader spectrum of the scientific community, so that emotionality does not impede objective assessment of experimental results" (Office of Technology Assessment, 1989).

It is in the spirit of the OTA report that this article is written. After Section 2, which offers an anecdotal account of the role of statisticians and statistics in parapsychology, the discussion turns to the more general question of replication of experimental results. Section 3 illustrates how replication has been (mis)interpreted by scientists in many fields. Returning to parapsychology in Section 4, a particular experimental regime called the "ganzfeld" is described, and an extended debate about the interpretation of the experimental results is discussed. Section

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

5 examines a meta-analysis of recent ganzfeld experiments designed to resolve the debate. Finally, Section 6 contains a brief account of meta-analyses that have been conducted in other areas of parapsychology, and conclusions are given in Section 7.

2. STATISTICS AND PARAPSYCHOLOGY

Parapsychology had its beginnings in the investigation of purported mediums and other anecdotal claims in the late 19th century. The Society for Psychical Research was founded in Britain in 1882, and its American counterpart was founded in Boston in 1884. While these organizations and their members were primarily involved with investigating anecdotal material, a few of the early researchers were already conducting "forced-choice" experiments such as card-guessing. (Forced-choice experiments are like multiple choice tests; on each trial the subject must guess from a small, known set of possibilities.) Notable among these was Nobel Laureate Charles Richet, who is generally credited with being the first to recognize that probability theory could be applied to card-guessing experiments (Rhine, 1977, p.26; Richet, 1884).

F.Y. Edgeworth, partly in response to what he considered to be incorrect analyses of these experiments, offered one of the earliest treatises on the statistical evaluation of forced-choice experiments in two articles published in the *Proceedings of the Society for Psychical Research* (Edgeworth, 1885, 1886). Unfortunately, as noted by Mauskopf and McVaugh (1979) in their historical account of the period, Edgeworth's papers were "perhaps too difficult for their immediate audience" (p. 105).

Edgeworth began his analysis by using Bayes Theorem to derive the formula for the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

5 examines a meta-analysis of recent ganzfeld experiments designed to resolve the debate. Finally, Section 6 contains a brief account of meta-analyses that have been conducted in other areas of parapsychology, and conclusions are given in Section 7.

2. STATISTICS AND PARAPSYCHOLOGY

Parapsychology had its beginnings in the investigation of purported mediums and other anecdotal claims in the late 19th century. The Society for Psychical Research was founded in Britain in 1882, and its American counterpart was founded in Boston in 1884. While these organizations and their members were primarily involved with investigating anecdotal material, a few of the early researchers were already conducting "forced-choice" experiments such as card-guessing. (Forced-choice experiments are like multiple choice tests; on each trial the subject must guess from a small, known set of possibilities.) Notable among these was Nobel Laureate Charles Richet, who is generally credited with being the first to recognize that probability theory could be applied to card-guessing experiments (Rhine, 1977, p.26; Richet, 1884).

F.Y. Edgeworth, partly in response to what he considered to be incorrect analyses of these experiments, offered one of the earliest treatises on the statistical evaluation of forced-choice experiments in two articles published in the *Proceedings of the Society for Psychical Research* (Edgeworth, 1885, 1886). Unfortunately, as noted by Mauskopf and McVaugh (1979) in their historical account of the period, Edgeworth's papers were "perhaps too difficult for their immediate audience" (p. 105).

Edgeworth began his analysis by using Bayes Theorem to derive the formula for the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

posterior probability that chance was operating, given the data. He then continued with an argument "savouring more of Bernoulli than Bayes" in which "it is consonant, I submit, to experience, to put ½ both for α and ß," i.e. for both the prior probability that chance alone was operating, and the prior probability that "there should have been some additional agency." He then reasoned (using a Taylor Series expansion of the posterior probability formula) that if there were a large probability of observing the data given that some additional agency was at work, and a small objective probability of the data under chance, then the latter (binomial) probability "may be taken as a rough measure of the sought *a posteriori* probability in favour of mere chance" (p. 195). Edgeworth concluded his article by applying his method to some data published previously in the same journal. He found the probability against chance to be .99996, which he said "may fairly be regarded as physical certainty" (p. 199). He concluded:

> "Such is the evidence which the calculus of probabilities affords as to the existence of an agency other than mere chance. The calculus is silent as to the nature of that agency -- whether it is more likely to be vulgar illusion or extraordinary law. That is a question to be decided, not by formulae and figures, but by general philosophy and common sense" (p. 199).

Both the statistical arguments and the experimental controls in these early experiments were somewhat loose. For example, Edgeworth treated as binomial an experiment in which one person chose a string of eight letters and another attempted to guess the string. Since it has long been understood that people are poor random number (or letter) generators, there is no statistical basis for analyzing such an experiment. Nonetheless, Edgeworth and his contemporaries set the stage for the use of controlled experiments with statistical evaluation in laboratory parapsychology.

One of the first American researchers to use statistical methods in parapsychology was

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

John Edgar Coover, who was the Thomas Welton Stanford Psychical Research Fellow, in the Psychology Department at Stanford University, from 1912 to 1937 (Dommeyer, 1975). In 1917 Coover published a large volume summarizing his work (Coover, 1917). Coover believed that his results were consistent with chance, but others have argued that Coover's definition of significance was too strict (Dommeyer, 1975). For example, in one evaluation of his telepathy experiments, Coover found a two-tailed p-value of .0062. He concluded "Since this value, then, lies within the field of chance deviation, although the probability of its occurrence by chance is fairly low, it cannot be accepted as a decisive indication of some cause beyond chance which operated in favor of success in guessing" (Coover, 1917, p. 82). On the next page he made it explicit that he would require a p-value of .0000221 to declare that something other than chance was operating.

It was during the summer of 1930, with the card-guessing experiments of J.B. Rhine at Duke University, that parapsychology began to take hold as a laboratory science. In fact, Rhine's laboratory still exists under the name of the Foundation for Research on the Nature of Man, housed at the edge of the Duke University campus.

It wasn't long after Rhine published his first book, *Extrasensory Perception* in 1934, that the attacks on his methodology began. Since his claims were wholly based on statistical analyses of his experiments, the statistical methods were closely scrutinized by critics anxious to find a plausible explanation for Rhine's positive results.

The most persistent critic was a psychologist from McGill University named Chester Kellogg (Mauskopf and McVaugh, 1979). Kellogg's main argument was that Rhine was using the binomial distribution (and normal approximation) on a series of trials that were not

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

independent. The experiments in question consisted of having a subject guess the order of a deck of 25 cards, with five each of five symbols, so technically Kellogg was correct.

By 1937 several mathematicians and statisticians had come to Rhine's aid. Mauskopf and McVaugh (1979) speculated that since statistics was itself a young discipline, "a number of statisticians were equally outraged by Kellogg, whose arguments they saw as discrediting *their* profession" (p. 258). The major technical work, which acknowledged that Kellogg's criticisms were accurate but did little to change the significance of the results, was conducted by Charles Stuart and Joseph A. Greenwood and published in the first volume of the *Journal of Parapsychology* (Stuart and Greenwood, 1937). Stuart, who had been an undergraduate in mathematics at Duke, was one of Rhine's early subjects, and continued to work with him as a researcher until Stuart's death in 1947. Greenwood was a Duke mathematician, who apparently converted to a statistician at the urging of Rhine.

Another prominent figure who was distressed with Kellogg's attack was E. V. Huntington, a mathematician at Harvard. After corresponding with Rhine, Huntington decided that, rather than further confuse the public with a technical reply to Kellogg's arguments, a simple statement should be made to the effect that the mathematical issues in Rhine's work had been resolved. Huntington must have successfully convinced his former student, Burton Camp of Wesleyan, that this was a wise approach. Camp was the 1937 President of IMS. When the annual meetings were held in December of 1937 (jointly with AMS and AAAS), Camp released a statement to the press that read:

> "Dr. Rhine's investigations have two aspects: experimental and statistical. On the experimental side mathematicians, of course, have nothing to say. On the statistical side, however, recent mathematical work has established the fact that, assuming that the experiments have been properly performed, the statistical

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

analysis is essentially valid. If the Rhine investigation is to be fairly attacked, it must be on other than mathematical grounds" (Camp, 1937).

One statistician who did emerge as a critic was William Feller. In a talk at the Duke Mathematical Seminar on April 24, 1940, Feller raised three criticisms to Rhine's work (Feller, 1940). They had been raised before by others (and continue to be raised even today). The first was that inadequate shuffling of the cards resulted in additional information from one series to the next. The second was what is now known as the "file-drawer effect," namely, that if one combines the results of published studies only, there is sure to be a bias in favor of successful studies. The third was that the results were enhanced by the use of optional stopping, i.e. by not specifying the number of trials in advance. All three of these criticisms were addressed in a rejoinder by Greenwood and Stuart (1940), but Feller was never convinced. Even in its third edition published in 1968, his book *An Introduction to Probability Theory and Its Applications* still contains his conclusion about Greenwood and Stuart: "Both their arithmetic and their experiments have a distinct tinge of the supernatural" (Feller, 1968, p. 407). In his discussion of Feller's position, Diaconis (1978) remarks, "I believe Feller was confused...he seemed to have decided the opposition was wrong and that was that."

Several statisticians have contributed to the literature in parapsychology to greater or lesser degrees. T.N.E. Greville devoted much of his professional life to developing statistical methods for parapsychology; Fisher (1924, 1929) addressed some specific problems in card-guessing experiments; Wilks (1965) described various statistical methods for parapsychology; Lindley (1957) presented a Bayesian analysis of some parapsychology data; and Diaconis (1978) pointed out some problems with certain experiments and presented a method for analyzing experiments when feedback is given.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Occasionally, attacks on parapsychology have taken the form of attacks on statistical inference in general, at least as it is applied to real data. Spencer-Brown (1957) attempted to show that true randomness is impossible, at least in finite sequences, and that this could be the explanation for the results in parapsychology. That argument re-emerged in a recent debate on the role of randomness in parapsychology, initiated by psychologist J. Barnard Gilmore (Gilmore, 1989; Utts, 1989a; Palmer, 1989; Gilmore, 1990; Palmer, 1990). Gilmore stated that "The agnostic statistician, advising on research in psi, should take account of the possible inappropriateness of classical inferential statistics" (1989, p.338). In his second paper, Gilmore reviewed several non-psi studies showing purportedly random systems that do not behave as they should under randomness (e.g. Iversen, Longcor, Mosteller, Gilbert, and Youtz, 1971; and Spencer-Brown, 1957). Gilmore concluded that "Anomalous data ...should not be found nearly so often if classical statistics offers a valid model of reality" (1990, p. 54), thus rejecting the use of classical statistical inference for real-world applications in general.

3. REPLICATION

Implicit and explicit in the literature on parapsychology is the assumption that in order to truly establish itself, the field needs to find a repeatable experiment. For example, Diaconis (1978) starts the summary of his article in *Science* with the words "In search of repeatable ESP experiments, modern investigators..." (p. 131). On October 28-29, 1983, the 32nd International Conference of the Parapsychology Foundation was held in San Antonio, Texas, to address "The Repeatability Problem in Parapsychology." The Conference Proceedings (Shapin and Coly, 1985) reflect the diverse views among parapsychologists on the nature of the problem. Honorton

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

(1985a) and Rao (1985), for example, both argued that strict replication is uncommon in *most* branches of science, and that parapsychology should not be singled out as unique in this regard. Other authors expressed disappointment in the lack of a single repeatable experiment in parapsychology, with titles such as "Unrepeatability: Parapsychology's Only Finding" (Blackmore, 1985), and "Research Strategies for Dealing with Unstable Phenomena" (Beloff, 1985).

It has never been clear, however, just exactly what would constitute acceptable evidence of a repeatable experiment. In the early days of investigation, the major critics "insisted that it would be sufficient for Rhine and Soal to convince them of ESP if a parapsychologist could perform successfully a single 'fraud-proof' experiment" (Hyman, 1985a, p. 71). However, as soon as well-designed experiments showing statistical significance emerged, the critics realized that a single experiment could be statistically significant just by chance. British psychologist C.E.M. Hansel quantified the new expectation, that the experiment should be repeated a few times, as follows:

> "If a result is significant at the .01 level and this result is not due to chance but to information reaching the subject, it may be expected that by making two further sets of trials the antichance odds of one hundred to one will be increased to around a million to one, thus enabling the effects of ESP -- or whatever is responsible for the original result -- to manifest itself to such an extent that there will be little doubt that the result is not due to chance" (Hansel, 1980, p.298).

In other words, three consecutive experiments at $p \leq .01$ would convince Hansel that something other than chance was at work.

This argument implies that if a particular experiment produces a statistically significant result, but subsequent replications fail to attain significance, then the original result was probably due to chance, or at least remains unconvincing. The problem with this line of reasoning is that

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

there is no consideration given to sample size or power. Only an experiment with extremely high power should be expected to be "successful" three times in succession.

It is perhaps a failure of the way statistics is taught that many scientists do not understand the importance of power in defining successful replication. To illustrate this point, psychologists Tversky and Kahnemann (1982) distributed a questionnaire to their colleagues at a professional meeting, with the question:

> "An investigator has reported a result that you consider implausible. He ran 15 subjects, and reported a significant value, $t = 2.46$. Another investigator has attempted to duplicate his procedure, and he obtained a nonsignificant value of t with the same number of subjects. The direction was the same in both sets of data. You are reviewing the literature. What is the highest value of t in the second set of data that you would describe as a failure to replicate?" (1982, p. 28).

In reporting their results, Tversky and Kahnemann stated:

> "The majority of our respondents regarded $t = 1.70$ as a failure to replicate. If the data of two such studies (t = 2.46 and $t = 1.70$) are pooled, the value of t for the combined data is about 3.00 (assuming equal variances). Thus, we are faced with a paradoxical state of affairs, in which the same data that would increase our confidence in the finding when viewed as part of the original study, shake our confidence when viewed as an independent study" (1982, p. 28).

At a recent presentation to the History and Philosophy of Science Seminar at the University of California at Davis, I asked the following question. Two scientists, Professors A and B, each have a theory they would like to demonstrate. Each plans to run a fixed number of Bernoulli trials and then test H_0: p = .25 versus H_a: p > .25. Professor A has access to large numbers of students each semester to use as subjects. In his first experiment he runs 100 subjects, and there are 33 successes ($p = .04$, one-tailed). Knowing the importance of replication, Professor A runs an additional 100 subjects as a second experiment. He finds 36 successes ($p = .009$, one-tailed).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Professor B only teaches small classes. Each quarter she runs an experiment on her students to test her theory. She carries out ten studies this way, with the following results:

n	Number of successes	one-tailed p-value
10	4	.22
15	6	.15
17	6	.23
25	8	.17
30	10	.20
40	13	.18
18	7	.14
10	5	.08
15	5	.31
20	7	.21

I asked the audience by a show of hands to indicate whether or not they felt the scientists had successfully demonstrated their theories. Professor A's theory received overwhelming support, with approximately 20 votes, while Professor B's theory received only one vote.

If you aggregate the results of the experiments for each Professor, you will notice that each conducted 200 trials, and Professor B actually demonstrated a *higher* level of success than Professor A, with 71 as opposed to 69 successful trials. The one-tailed *p*-values for the combined trials are .0017 for Professor A and .0006 for Professor B.

To address the question of replication more explicitly, I also posed the following scenario. In December of 1987 it was decided to prematurely terminate a study on the effects of aspirin in reducing heart attacks because the data were so convincing (See e.g. Greenhouse and Greenhouse, 1988; Rosenthal, 1990a). The physician-subjects had been randomly assigned to take aspirin or a placebo. There were 104 heart attacks among the 11,037 subjects in the aspirin group, and 189 heart attacks among the 11,034 subjects in the placebo group (chi-square

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

= 25.01, $p<.00001$).

After showing the results of that study, I presented the audience with two hypothetical experiments conducted to try to replicate the original result, with outcomes as follows:

	REPLICATION #1			REPLICATION #2	
	Heart Attack			Heart Attack	
	Yes	No		Yes	No
Aspirin	11	1156	Aspirin	20	2314
Placebo	19	1090	Placebo	48	2170
Chi-square = 2.596, $p=.11$			Chi-square = 13.206, $p=.0003$		

I asked the audience to indicate which one they thought was a more successful replication. The audience chose the second one, as would most journal editors, because of the "significant p-value". In fact, the *first* replication has almost exactly the same proportion of heart attacks in the two groups as the original study, and is thus a very close replication of that result. The second replication has very *different* proportions, and in fact the relative risk from the second study is not even contained in a 95% confidence interval for relative risk from the original study. The *magnitude* of the effect has been much more closely matched by the "non-significant" replication.

Fortunately, psychologists are beginning to notice that replication is not as straightforward as they were originally led to believe. A special issue of the *Journal of Social Behavior and Personality* was entirely devoted to the question of replication (Neuliep, 1990). In one of the articles, Rosenthal cautioned his colleagues: "Given the levels of statistical power at which we normally operate, we have no right to expect the proportion of significant results

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

that we typically do expect, even if in nature there is a very real and very important effect" (Rosenthal, 1990b, p.16).

Jacob Cohen, in his insightful article titled "Things I Have Learned (So Far)," identified another misconception common among social scientists: "Despite widespread misconceptions to the contrary, the rejection of a given null hypothesis gives us no basis for estimating the probability that a replication of the research will again result in rejecting that null hypothesis" (Cohen, 1990, p.1307).

Cohen and Rosenthal both advocate the use of effect sizes as opposed to significance levels when defining the strength of an experimental effect. In general, effect sizes measure the amount by which the data deviate from the null hypothesis in terms of standardized units. For instance, the effect size for a two-sample *t*-test is usually defined to be the difference in the two means, divided by the standard deviation for the control group. This measure can be compared across studies without the dependence on sample size inherent in significance levels. (Of course there will still be variability in the sample effect sizes, decreasing as a function of sample size.) Comparison of effect sizes across studies is one of the major components of meta-analysis.

Similar arguments have recently been made in the medical literature. For example, Gardner and Altman (1986) stated that the use of *p*-values "to define two alternative outcomes - significant and not significant - is not helpful and encourages lazy thinking" (p. 746). They advocated the use of confidence intervals instead.

As discussed in the next section, the arguments used to conclude that parapsychology has failed to demonstrate a replicable effect hinge on these misconceptions of replication and failure to examine power. A more appropriate analysis would compare the effect sizes for similar

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

experiments across experimenters and across time to see if there have been consistent effects of the same magnitude. Rosenthal also advocates this view of replication:

> "The traditional view of replication focuses on significance level as the relevant summary statistic of a study and evaluates the success of a replication in a dichotomous fashion. The newer, more useful view of replication focuses on effect size as the more important summary statistic of a study and evaluates the success of a replication not in a dichotomous but in a continuous fashion" (Rosenthal, 1990b, p. 28).

The dichotomous view of replication has been used throughout the history of parapsychology, by both parapsychologists and critics (Utts, 1988). For example, the National Academy of Sciences Report critically evaluated "significant" experiments, but entirely ignored "nonsignificant" experiments.

In the next three sections we will examine some of the results in parapsychology using the broader, more appropriate definition of replication. In doing so, we will show that the results are far more interesting than the critics would have us believe.

4. THE GANZFELD DEBATE IN PARAPSYCHOLOGY

An extensive debate took place in the mid-1980's between a parapsychologist and critic, questioning whether or not a particular body of parapsychological data had demonstrated psi abilities. The experiments in question were all conducted using the ganzfeld setting (described below). Several authors were invited to write commentaries on the debate. As a result, this data base has been more thoroughly analyzed by both critics and proponents than any other, and provides a good source for studying replication in parapsychology.

The debate concluded with a detailed series of recommendations for further experiments, and left open the question of whether or not psi abilities had been demonstrated. A new series

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

of experiments that followed the recommendations were conducted over the next few years. The results of the new experiments will be presented in Section 5.

4.1 Free-response Experiments

Recent experiments in parapsychology tend to use more complex target material than the cards and dice used in the early investigations, partially to alleviate boredom on the part of the subjects and partially because they are thought to "more nearly resemble the conditions of spontaneous psi occurrences" (Burdick and Kelly, 1977, p. 109). These experiments fall under the general heading of "free-response" experiments, because the subject is asked to give a verbal or written description of the target, rather than being forced to make a choice from a small discrete set of possibilities. Various types of target material have been used, including pictures, short segments of movies on video tapes, actual locations, and small objects.

Despite the more complex target material, the statistical methods used to analyze these experiments are similar to those for forced-choice experiments. A typical experiment proceeds as follows. Before conducting any trials, a large pool of potential targets is assembled, usually in packets of four. Similarity of targets within a packet is kept to a minimum, for reasons made clear below. At the start of an experimental session, after the subject is sequestered in an isolated room, a target is selected at random from the pool. A sender is placed in another room with the target. The subject is asked to provide a verbal or written description of what he or she thinks is in the target, knowing only that it is a photograph, an object, etc.

After the subject's description has been recorded and secured against the potential for later alteration, a judge (who may or may not be the subject) is given a copy of the subject's description and the four possible targets that were in the packet with the correct target. A

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

properly conducted experiment either uses video tapes or has two identical sets of target material and uses the duplicate set for this part of the process, to ensure that clues such as fingerprints don't give away the answer. Based on the subject's description, and of course on a blind basis, the judge is asked to either rank the four choices from most to least likely to have been the target, or to select the one from the four that seems to best match the subject's description. If ranks are used, the statistical analysis proceeds by summing the ranks over a series of trials and comparing the sum to what would be expected by chance. If the selection method is used, a "direct hit" occurs if the correct target is chosen, and the number of direct hits over a series of trials is compared to the number expected in a binomial experiment with $p = .25$.

Note that the subjects' *responses* cannot be considered to be "random" in any sense, so probability assessments are based on the random selection of the target and decoys. In a correctly designed experiment, the probability of a direct hit by chance is .25 on each trial, regardless of the response, and the trials are independent. These and other issues related to analyzing free-response experiments are discussed by Utts (1989b).

4.2 The Psi Ganzfeld Experiments

The ganzfeld procedure is a particular kind of free-response experiment utilizing a perceptual isolation technique originally developed by Gestalt psychologists for other purposes. Evidence from spontaneous case studies and experimental work had led parapsychologists to a model proposing that psychic functioning may be masked by sensory input and by inattention to internal states (Honorton, 1977). The ganzfeld procedure was specifically designed to test whether or not reduction of external "noise" would enhance psi performance.

In these experiments, the subject is placed in a comfortable reclining chair in an

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

acoustically shielded room. To create a mild form of sensory deprivation, the subject wears headphones through which white noise is played, and stares into a constant field of red light. This is achieved by taping halved translucent ping-pong balls over the eyes and then illuminating the room with red light. In the psi ganzfeld experiments, the subject speaks into a microphone and attempts to describe the target material being observed by the sender in a distant room.

At the 1982 Annual Meeting of the Parapsychological Association, a debate took place over the degree to which the results of the psi ganzfeld experiments constituted evidence of psi abilities. Psychologist and critic Ray Hyman and parapsychologist Charles Honorton each analyzed the results of all known psi ganzfeld experiments to date, and reached strikingly different conclusions. The debate continued with the publication of their arguments in separate articles in the March 1985 issue of the *Journal of Parapsychology*. Finally, in the December 1986 issue of the *Journal of Parapsychology*, Hyman and Honorton wrote a joint article in which they highlighted their agreements and disagreements, and outlined detailed criteria for future experiments. That same issue contained commentaries on the debate by ten other authors.

The data base analyzed by Hyman and Honorton consisted of results taken from 34 reports written by a total of 47 authors. Honorton counted 42 separate experiments described in the reports, of which 28 reported enough information to determine the number of direct hits achieved. Twenty three of the studies (55%) were classified by Honorton as having achieved statistical significance at .05.

4.3 The Vote-Counting Debate

Vote-counting is the term commonly used for the technique of drawing inferences about an experimental effect by counting the number of significant versus non-significant studies of

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the effect. Hedges and Olkin (1985) give a detailed analysis of the inadequacy of this method, showing that it is more and more likely to make the wrong decision as the number of studies increases. While Hyman acknowledged that "vote-counting raises many problems (Hyman, 1985b, p.8)," he nonetheless spent half of his critique of the ganzfeld studies showing why Honorton's count of 55% was wrong.

Hyman's first complaint was that several of the studies contained multiple conditions, each of which should be considered as a separate study. Using this definition he counted 80 studies (thus further reducing the sample sizes of the individual studies), of which 25 (31%) were "successful." Honorton's response to this was to invite readers to examine the studies and decide for themselves if the varying conditions constituted separate experiments.

Hyman next postulated that there was selection bias, so that significant studies were more likely to be reported. He raised some important issues about how pilot studies may be terminated and not reported if they don't show significant results, or may at least be subject to optional stopping, allowing the experimenter to determine the number of trials. He also presented a chi-square analysis that "suggests a tendency to report studies with a small sample only if they have significant results" (Hyman, 1985b, p.14). but I have questioned his analysis elsewhere (Utts, 1986, p. 397).

Honorton refuted Hyman's argument with four rejoinders (Honorton, 1985b, p.66). In addition to reinterpreting Hyman's chi-square analysis, Honorton pointed out that the Parapsychological Association has an official policy encouraging the publication of non-significant results in its journals and proceedings, that a large number of reported ganzfeld studies did not achieve statistical significance, and that there would have to be 15 studies in the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

"file-drawer" for every one reported to cancel out the observed significant results.

The remainder of Hyman's vote-counting analysis consisted of showing that the effective error rate for each study was actually much higher than the nominal 5%. For example, each study could have been analyzed using the direct hit measure, the sum of ranks measure, or one of two other measures used for free-response analyses. Hyman carried out a simulation study that showed the true error rate would be .22 if "significance" was defined by requiring at least one of these four measures to achieve the .05 level. He suggested several other ways in which multiple testing could occur, and concluded that the effective error rate in each experiment was not the nominal .05, but rather was probably close to the 31% he had determined to be the actual success rate in his vote-count.

Honorton acknowledged that there was a multiple testing problem, but he had a two-fold response. First, he applied a Bonferroni correction and found that the number of significant studies (using his definition of a study) only dropped from 55% to 45%. Next, he proposed that a uniform index of success be applied to all studies. He used the number of direct hits, since it was by far the most commonly reported measure and was the measure used in the first published psi ganzfeld study. He then conducted a detailed analysis of the 28 studies reporting direct hits and found that 43% were significant at .05 on that measure alone. Further, he showed that significant effects were reported by six of the 10 independent investigators, and thus were not due to just one or two investigators or laboratories. He also noted that success rates were very similar for reports published in refereed journals and those published in unrefereed monographs and abstracts.

While Hyman's arguments identified issues such as selective reporting and optional

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

stopping that should be considered in any meta-analysis, the dependence of significance levels on sample size makes the vote-counting technique almost useless for assessing the magnitude of the effect. Consider for example the 24 studies where the direct hit measure was reported and the chance probability of a direct hit was .25, the most common type of study in the data base. (There were 4 direct hit studies with other chance probabilities and 14 that did not report direct hits.) Of the 24 studies, 13 (54%) were "nonsignificant" at $\alpha = .05$, one-tailed. But if the 367 trials in these "failed replications" are combined, there are 106 direct hits, $z = 1.66$, and $p = .0485$, one tailed. This is reminiscent of the dilemma of Professor B in Section 3.

Power is typically very low for these studies. The median sample size for the studies reporting direct hits was 28. If there is a real effect and it increases the success probability from the chance .25 to an actual .33 (a value whose rationale will be made clear below), the power for a study with 28 trials is only .181 (Utts, 1986). It should be no surprise that there is a "repeatability" problem in parapsychology.

4.4 Flaw Analysis and Future Recommendations

The second half of Hyman's paper consisted of a "Meta-Analysis of Flaws and Successful Outcomes" (1985b, p. 30), designed to explore whether or not various measures of success were related to specific flaws in the experiments. While many critics have argued that the results in parapsychology can be explained by experimental flaws, Hyman's analysis was the first to attempt to quantify the relationship between flaws and significant results.

Hyman identified 12 potential flaws in the ganzfeld experiments, such as inadequate randomization, multiple tests used without adjusting the significance level (thus inflating the significance level from the nominal 5%), and failure to use a duplicate set of targets for the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

judging process (thus allowing possible clues such as fingerprints). Using cluster and factor analyses, the 12 binary flaw variables were combined into three new variables, which Hyman named General Security, Statistics and Controls.

Several analyses were then conducted. The one reported with the most detail is a factor analysis utilizing 17 variables for each of 36 studies. Four factors emerged from the analysis. From these, Hyman concluded that security had increased over the years, that the significance level tended to be inflated the most for the most complex studies, and that both effect size and level of significance were correlated with the existence of flaws.

Following his factor analysis, Hyman picked the three flaws that seemed to be most highly correlated with success, which were inadequate attention to both randomization and documentation, and the potential for ordinary communication between the sender and receiver. A regression equation was then computed using each of the three flaws as dummy variables, and the effect size for the experiment as the dependent variable. From this equation, Hyman concluded that a study without these three flaws would be predicted to have a hit rate of 27%. He concluded that this is "well within the statistical neighborhood of the 25% chance rate" (ibid, p. 37), and thus "the ganzfeld psi data base, despite initial impressions, is inadequate either to support the contention of a repeatable study or to demonstrate the reality of psi" (ibid p. 38).

Honorton discounted both Hyman's flaw classification and his analysis. He did not deny that flaws existed, but objected that Hyman's analysis was faulty and impossible to interpret. Honorton asked psychometrician David Saunders to write an Appendix to his article, evaluating Hyman's analysis. Saunders first criticized Hyman's use of a factor analysis with 17 variables (many of which were dichotomous) and only 36 cases, and concluded that "the entire analysis

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

is meaningless" (Saunders, 1985, p.87). He then noted that Hyman's choice of the three flaws to include in his regression analysis constituted a clear case of multiple analysis, since there were 84 possible sets of three that could have been selected (out of nine potential flaws), and Hyman chose the set most highly correlated with effect size. Again, Saunders concluded that "any interpretation drawn from [the regression analysis] must be regarded as meaningless" (ibid, p. 88).

Hyman's results were also contradicted by Harris and Rosenthal (1988b) in an analysis requested by Hyman in his capacity as Chair of the National Academy of Sciences' Subcommittee on Parapsychology. Using Hyman's flaw classifications and a multivariate analysis, Harris and Rosenthal concluded that "Our analysis of the effects of flaws on study outcome lends no support to the hypothesis that ganzfeld research results are a significant function of the set of flaw variables" (1988b, p. 3).

Hyman and Honorton were in the process of preparing papers for a second round of debate when they were invited to lunch together at the 1986 Meeting of the Parapsychological Association. They discovered that they were in general agreement on several major issues, and decided to coauthor a "Joint Communique" (Hyman and Honorton, 1986). It is clear from their paper that they both thought it was more important to set the stage for future experimentation than to continue the technical arguments over the current data base. In the abstract to their paper they wrote:

> "We agree that there is an overall significant effect in this data base that cannot reasonably be explained by selective reporting or multiple analysis. We continue to differ over the degree to which the effect constitutes evidence for psi, but we agree that the final verdict awaits the outcome of future experiments conducted by a broader range of investigators and according to more stringent standards" (Ibid, p. 351).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The paper then outlined what these standards should be. They included controls against any kind of sensory leakage, thorough testing and documentation of randomization methods used, better reporting of judging and feedback protocols, control for multiple analyses, and advance specification of number of trials and type of experiment. Indeed, any area of research could benefit from such a careful list of procedural recommendations.

4.5 Rosenthal's Meta-Analysis

The same issue of the *Journal of Parapsychology* in which the Joint Communique appeared also carried commentaries on the debate by 10 separate authors. In his commentary, psychologist Robert Rosenthal, one of the pioneers of meta-analysis in psychology, summarized the aspects of Hyman's and Honorton's work that would typically be included in a meta-analysis (Rosenthal, 1986). It is worth reviewing Rosenthal's results so that they can be used as a basis of comparison for the more recent psi ganzfeld studies reported in Section 5.

Rosenthal, like Hyman and Honorton, focused only on the 28 studies for which direct hits were known. He chose to use an effect size measure called Cohen's h, which is the difference between the arcsin transformed proportions of direct hits that were observed and expected:

$$h=2\times(\arcsin\sqrt{\hat{p}}-\arcsin\sqrt{p})$$

One advantage of this measure over the difference in raw proportions is that can be used to compare experiments with different chance hit rates.

If the observed and expected numbers of hits were identical, the effect size would be zero. Of the 28 studies, 23 (82%) had effect sizes greater than zero, with a median effect size

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

of .32 and a mean of .28. These correspond to direct hit rates of .40 and .38 respectively, when .25 is expected by chance. A 95% confidence interval for the true effect size is from .11 to .45, corresponding to direct hit rates of from .30 to .46 when chance is .25.

A common technique in meta-analysis is to calculate a "combined *z*," found by summing the individual *z* scores and dividing by the square root of the number of studies. The result should have a standard normal distribution if each *z* score has a standard normal distribution. For the ganzfeld studies, Rosenthal reported a combined *z* of 6.60 with a *p*-value of 3.37×10^{-11}. He also reiterated Honorton's file-drawer assessment by calculating that there would have to be 423 studies unreported to negate the significant effect in the 28 direct hit studies.

Finally, Rosenthal acknowledged that because of the flaws in the data base and the potential for at least a small file drawer effect, the true average effect size was probably closer to .18 than .28. He concluded, "Thus, when the accuracy rate expected under the null is 1/4, we might estimate the obtained accuracy rate to be about 1/3" (Ibid, p. 333). This is the value used for the earlier power calculation.

It is worth mentioning that Rosenthal was commissioned by the National Academy of Sciences to prepare a background paper to accompany its 1988 report on parapsychology. That paper (Harris and Rosenthal, 1988a) contained much of the same analysis as his commentary summarized above. Ironically, the discussion of the ganzfeld work in the National Academy Report focused on Hyman's 1985 analysis, but never mentioned the work it had commissioned Rosenthal to perform, which contradicted the final conclusion in the report.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

5. A META-ANALYSIS OF RECENT GANZFELD EXPERIMENTS

After the initial exchange with Hyman at the 1982 Parapsychological Association Meeting, Honorton and his colleagues developed an automated ganzfeld experiment, that was designed to eliminate the methodological flaws identified by Hyman. The execution and reporting of the experiments followed the detailed guidelines agreed upon by Hyman and Honorton.

Using this "autoganzfeld" experiment, eleven experimental series were conducted by eight experimenters between February 1983 and September 1989, when the equipment had to be dismantled due to lack of funding. In this section the results of these experiments are summarized and compared to the earlier ganzfeld studies. Much of the information is derived from Honorton et al (1990).

5.1 The Automated Ganzfeld Procedure

Like earlier ganzfeld studies, the "autoganzfeld" experiments require four participants. The first is the Receiver (R), who attempts to identify the target material being observed by the Sender (S). The Experimenter (E) prepares R for the task, elicits the response from R, and supervises R's judging of the response against the four potential targets. (Judging is double-blind; E does not know which is the correct target.) The fourth participant is the lab assistant (LA) whose only task is to instruct the computer to randomly select the target. No one involved in the experiment knows the identity of the target.

Both R and S are sequestered in sound-isolated, electrically shielded rooms. R is prepared as in earlier ganzfeld studies, with white noise and a field of red light. In a

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

non-adjacent room, S watches the target material on a television and can hear R's target description ("mentation") as it is being given. The mentation is also tape-recorded.

The judging process takes place immediately after the 30 minute sending period. On a TV monitor in the isolated room, R views the four choices from the target pack that contains the actual target. R is asked to rate each one according to how closely it matches the ganzfeld mentation. The ratings are converted to ranks, and if the correct target is ranked first, a direct hit is scored. The entire process is automatically recorded by the computer. The computer then displays the correct choice to R as feedback.

There were 160 pre-selected targets, used with replacement, in ten of the eleven series. They were arranged in packets of 4, and the decoys for a given target were always the remaining three in the same set. Thus, even if a particular target in a set were consistently favored by R's, the probability of a direct hit under the null hypothesis would remain at 1/4. Popular targets should be no more likely to be selected by the computer's random number generator than any of the others in the set. The selection of the target by the computer is the only source of randomness in these experiments. This is an important point, and one that is often misunderstood. (See Utts, 1989b for elucidation.)

Eighty of the targets were "dynamic," consisting of scenes from movies, documentaries and cartoons; and 80 were "static", consisting of photographs, art prints, and advertisements. The four targets within each set were all of the same type. Earlier studies indicated that dynamic targets were more likely to produce successful results, and one of the goals of the new experiments was to test that theory.

The randomization procedure used to select the target and the order of presentation for

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

judging was thoroughly tested before and during the experiments. A detailed description is given by Honorton et al (1990, p. 118-120).

Three of the eleven series were pilot series, five were formal series with novice receivers, and three were formal series with experienced receivers. The last series with experienced receivers was the only one that did not use the 160 targets. Instead, it used only one set of four dynamic targets in which one target had previously received several first place ranks, and one had never received a first place rank. The receivers, none of whom had had prior exposure to that target pack, were not aware that only one target pack was being used. They each contributed one session only to the series. This will be called the "special series" in what follows.

Except for two of the pilot series, numbers of trials were planned in advance for each series. Unfortunately, three of the formal series were not yet completed when the funding ran out, including the special series, and one pilot study with advance planning was terminated early when the experimenter relocated. There were no unreported trials during the six year period under review, so there was no "file-drawer".

Overall, there were 183 R's who contributed only one trial and 58 who contributed more than one, for a total of 241 participants and 355 trials. Only twenty three R's had previously participated in ganzfeld experiments and 194 R's (81%) had never participated in any parapsychological research.

5.2 Results

While acknowledging that no probabilistic conclusions can be drawn from qualitative data, Honorton et al (1990), included several examples of session excerpts that R's identified as

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

providing the basis for their target rating. To give a flavor for the dream-like quality of the mentation and the amount of information that can be lost by only assigning a rank, the first example is reproduced here. The target was a painting by Salvador Dali called "Christ Crucified." The correct target received a first place rank. The part of the mentation R used to make this assessment read:

> "... I think of guides, like spirit guides, leading me and I come into a court with a king. It's quiet.... It's like heaven. The king is something like Jesus. Woman. Now I'm just sort of summersaulting through heaven.... Brooding.... Aztecs, the Sun God.... High priest.... Fear.... Graves. Woman. Prayer.... Funeral.... Dark. Death.... Souls.... Ten Commandments. Moses...." (Ibid, p. 120).

Over all eleven series there were 122 direct hits in the 355 trials, for a hit rate of 34.4% (exact binomial p-value = .00005) when 25% were expected by chance. Cohen's h is .20, and a 95% confidence interval for the overall hit rate is from .30 to .39. This calculation assumes, of course, that the probability of a direct hit is constant and independent across trials, an assumption that may be questionable except under the null hypothesis of no psi abilities.

Honorton et al also calculated effect sizes for each of the eleven series and each of the eight experimenters. All but one of the series (the first novice series) had positive effect sizes, as did all of the experimenters.

The special series with experienced R's had an exceptionally high effect size with h = .81, corresponding to 16 direct hits out of 25 trials (64%), but the remaining series and the experimenters had relatively homogeneous effect sizes given the amount of variability expected by chance. If the special series is removed, the overall hit rate is 32.1%, h = .16. Thus, the positive effects are not due to just one series or one experimenter.

Seventy one of the 218 trials contributed by novices were direct hits (32.5%, h = .17),

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

compared with 51 hits in the 137 trials by those with prior ganzfeld experience (37%, h = .26). The hit rates and effect sizes were 31% (h = .14) for the combined pilot series, 32.5% (h = .17) for the combined formal novice series, and 41.5% (h = .35) for the combined experienced series. The last figure drops to 31.6% if the outlier series is removed. Finally, without the outlier series the hit rate for the combined series where all of the planned trials were completed was 31.2% (h = .14) while it was 35% (h = .22) for the combined series that were terminated early. Thus, optional stopping cannot account for the positive effect.

There were two interesting comparisons that had been suggested by earlier work and were preplanned in these experiments. The first was to compare results for trials with dynamic targets with those for static targets. In the 190 dynamic target sessions there were 77 direct hits (40%, h = .32) and for the static targets there were 45 hits in 165 trials (27%, h = .05), thus indicating that dynamic targets produced far more successful results.

The second comparison of interest was whether or not the sender was a friend of the receiver. This was a choice the receiver could make. If he or she did not bring a friend, a lab member acted as sender. There were 211 trials with friends as senders (some of whom were also lab staff), resulting in 76 direct hits (36%, h = .24). Four trials used no sender. The remaining 140 trials used non-friend lab staff as senders and resulted in 46 direct hits (33%, h = .18). Thus, trials with friends as senders were slightly more successful than those without.

Consonant with the definition of replication based on consistent effect sizes, it is informative to compare the autoganzfeld experiments with the direct hit studies in the previous data base. The overall success rates are extremely similar. The overall direct hit rate was 34.4% for the autoganzfeld studies and was 38% for the comparable direct hit studies in the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

earlier meta-analysis. Rosenthal's (1986) adjustment for flaws had placed a more conservative estimate at 33%, very close to the observed 34.4% in the new studies.

One limitation of this work is that the autoganzfeld studies, while conducted by eight experimenters, all used the same equipment in the same laboratory. Unfortunately, the level of funding available in parapsychology and the cost in time and equipment to conduct proper experiments make it difficult to amass large amounts of data across laboratories. Another autoganzfeld laboratory is currently being constructed at the University of Edinburgh in Scotland, so interlaboratory comparisons may be possible in the near future.

Based on the effect size observed to date, large samples are needed to achieve reasonable power. If there is a constant effect across all trials, resulting in 33% direct hits when 25% are expected by chance, to achieve a one tailed significance level of .05 with 95% probability would require 345 sessions.

We end this section by returning to the aspirin and heart attack example in Section 3, and expanding a comparison noted by Atkinson et al (1990, p. 237). Computing the equivalent of Cohen's h for comparing observed heart attack rates in the aspirin and placebo groups results in $h = .068$. Thus, the effect size observed in the ganzfeld data base is triple the much-publicized effect of aspirin on heart attacks.

6. OTHER META-ANALYSES IN PARAPSYCHOLOGY

Four additional meta-analyses have been conducted in various areas of parapsychology since the original ganzfeld meta-analyses were reported. Three of the four analyses focused on evidence of psi abilities, while the fourth examined the relationship between extraversion and

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

psychic functioning. In this section, each of the four analyses will be briefly summarized.

There are only a handful of English-language journals and proceedings in parapsychology, so retrieval of the relevant studies in each of the four cases was simple to accomplish by searching those sources in detail and by searching other bibliographic data bases for keywords.

Each analysis included an overall summary, an analysis of the quality of the studies versus the size of the effect, and a "file-drawer" analysis to determine the possible number of unreported studies. Three of the four also contained comparisons across various conditions.

6.1 Forced-choice Precognition Experiments

Honorton and Ferrari (1989) analyzed forced-choice experiments conducted from 1935 to 1987, in which the target material was randomly selected *after* the subject had attempted to predict what it would be. The time delay in selecting the target ranged from under a second to one year. Target material included items as diverse as ESP cards and automated random number generators. Two investigators, S.G. Soal and Walter J. Levy, were not included because some of their work has been suspected to be fraudulent.

Overall Results. There were 309 studies reported by 62 senior authors, including more than 50,000 subjects and nearly two million individual trials. Honorton and Ferrari used $z/\sqrt{n}$ as the measure of effect size (*ES*) for each study, where n was the number of Bernoulli trials in the study. They reported a mean *ES* of 0.020, and a mean z-score of 0.65 over all studies. They also reported a combined z of 11.41, $p = 6.3 \times 10^{-25}$. Thirty percent (92) of the studies were statistically significant at $\alpha = .05$. The mean *ES* per investigator was 0.033, and the significant results were not due to just a few investigators.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Quality. Eight dichotomous quality measures were assigned to each study, resulting in possible scores from zero for the lowest quality, to eight for the highest. They included features such as adequate randomization, preplanned analysis, and automated recording of the results. The correlation between study quality and effect size was 0.081, indicating a slight tendency for higher quality studies to be more successful, contrary to claims by critics that the opposite would be true. There was a clear relationship between quality and year of publication, presumably because over the years experimenters in parapsychology have responded to suggestions from critics for improving their methodology.

File-drawer. Following Rosenthal (1984), the authors calculated the "fail-safe *N*" indicating the number of unreported studies that would have to be sitting in file-drawers in order to negate the significant effect. They found $N = 14{,}268$, or a ratio of 46 unreported studies for each one reported. They also followed a suggestion by Dawes et al (1984) and computed the mean z for all studies with $z > 1.65$. If such studies were a random sample from the upper 5% tail of a N(0,1) distribution, the mean z would be 2.06. In this case it was 3.61. They concluded that selective reporting could not explain these results.

Comparisons. Four variables were identified that appeared to have a systematic relationship to study outcome. The first was that the 25 studies using subjects selected on the basis of good past performance were more successful than the 223 using unselected subjects, with mean effect sizes of .051 and .008, respectively. Second, the 97 studies testing subjects individually were more successful than the 105 studies that used group testing; mean effect sizes were .021 and .004, respectively. Timing of feedback was the third moderating variable, but information was only available for 104 studies. The 15 studies that never told the subjects what

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the targets were had a mean effect size of -.001. Feedback after each trial produced the best results, the mean *ES* for the 47 studies was .035. Feedback after each set of trials resulted in mean *ES* of .023 (21 studies), while delayed feedback (also 21 studies) yielded a mean *ES* of only .009. There is a clear ordering; as the gap between time of feedback and time of the actual guesses decreased, effect sizes increased.

The fourth variable was the time interval between the subject's guess and the actual target selection, available for 144 studies. The best results were for the 31 studies that generated targets less than a second after the guess (mean *ES* = .045), while the worst were for the 7 studies that delayed target selection by at least a month (mean *ES* = .001). The mean effect sizes showed a clear trend, decreasing in order as the time interval increased from minutes to hours to days to weeks to months.

6.2 Attempts to Influence Random Physical Systems

Radin and Nelson (1989) examined studies designed to test the hypothesis that "The statistical output of an electronic RNG [random number generator] is correlated with observer intention in accordance with prespecified instructions" (p. 1502). These experiments typically involve RNGs based on radioactive decay, electronic noise, or pseudorandom number sequences seeded with true random sources. Usually the subject is instructed to try to influence the results of a string of binary trials by mental intention alone. A typical protocol would ask a subject to press a button (thus starting the collection of a fixed-length sequence of bits), and then try to influence the random source to produce more zeroes or more ones. A run might consist of three successive button presses, one each in which the desired result was more zeroes or more ones, and one as a control with no conscious intention. A z score would then be computed for each

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

button press.

The 832 studies in the analysis were conducted from 1959 to 1987, and included 235 "control" studies in which the output of the RNGs were recorded but there was no conscious intention involved. These were usually conducted before and during the experimental series, as tests of the RNGs.

Results. The effect size measure used was again $z / \sqrt{n}$, where z was positive if more bits of the specified type were achieved. The mean effect size for control studies was not significantly different from zero (-1.0×10^{-5}). The mean effect size for the experimental studies was also very small, 3.2×10^{-4}, but it was significantly higher than the mean *ES* for the control studies ($z = 4.1$).

Quality. Sixteen quality measures were defined and assigned to each study, under the four general categories of procedures, statistics, data, and the RNG device. A score of 16 reflected the highest quality. The authors regressed mean effect size on mean quality for each investigator, and found a slope of 2.5×10^{-5} with standard error of 3.2×10^{-5}, indicating little relationship between quality and outcome. They also calculated a weighted mean effect size, using quality scores as weights, and found that it was very similar to the unweighted mean *ES*. They concluded that "differences in methodological quality are not significant predictors of effect size" (p. 1507).

File-drawer. Radin and Nelson used several methods for estimating the number of unreported studies (p. 1508-10). Their estimates ranged from 200 to 1000 based on models assuming that all significant studies were reported. They also calculated the fail-safe *N* to be 54,000.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

6.3 Attempts to Influence Dice

Radin and Ferrari (1991) examined 148 studies, published from 1935 to 1987, designed to test whether or not consciousness can influence the results of tossing dice. The also found 31 "control" studies in which no conscious intention was involved.

Results. The effect size measure used was $z / \sqrt{n}$, where z was based on the number of throws in which the die landed with the desired face (or faces) up, in n throws. The weighted mean *ES* for the experimental studies was 0.0122 with a standard error of 0.00062; for the control studies the mean and standard error were 0.00093 and 0.00255, respectively. Weights for each study were determined by quality, giving more weight to high quality studies. Combined z scores for the experimental and control studies were reported by Radin and Ferrari to be 18.2 and 0.18, respectively.

Quality. Eleven dichotomous quality measures were assigned, ranging from automated recording to whether or not control studies were interspersed with the experimental studies. The final quality score for each study combined these with information on method of tossing the dice, and with source of subject (defined below). A regression of quality score versus effect size resulted in a slope of -.002, with a standard error of .0011. However, when effect sizes were weighted by sample size there was a significant relationship between quality and effect size, leading Radin and Ferrari to conclude that higher quality studies produced lower weighted effect sizes.

File-drawer. Radin and Ferrari calculated Rosenthal's fail-safe *N* for this analysis to be 17,974. Using the assumption that all significant studies were reported, they estimated the number of unreported studies to be 1,152. As a final assessment, they compared studies

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

published before and after 1975, when the *Journal of Parapsychology* adopted an official policy of publishing nonsignificant results. They concluded, based on that analysis, that more nonsignificant studies were published after 1975, and thus "We must consider the overall (1935-1987) data base as suspect with respect to the filedrawer problem."

Comparisons. Radin and Ferrari noted that there was bias in both the experimental and control studies across die face. Six was the face most likely to come up, consistent with the observation that it has the least mass. Therefore, they examined results for the subset of 69 studies in which targets were evenly balanced among the six faces. They still found a significant effect, with mean and standard error for effect size of 8.6×10^{-3} and 1.1×10^{-3}, respectively. The combined z was 7.617 for these studies.

They also compared effect sizes across types of subjects used in the studies, categorizing them as unselected, experimenter and other subjects, experimenter as sole subject, and specially selected subjects. Like Honorton and Ferrari (1989), they found the highest mean *ES* for studies with selected subjects; it was approximately .02, more than twice that for unselected subjects.

6.4 Extraversion and ESP Performance

Honorton, Ferrari and Bem (1990) conducted a meta-analysis to examine the relationship between scores on tests of extraversion and scores on psi-related tasks. They found 60 studies by 17 investigators, conducted from 1945 to 1983.

Results. The effect size measure used for this analysis was the correlation between each subject's extraversion score and ESP score. A variety of measures had been used for both scores across studies, so various correlation coefficients were used. Nonetheless, a stem and leaf diagram of the correlations showed an approximate bell shape with mean and standard

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

deviation of .19 and .26, respectively, and with an additional outlier at $r = .91$. Honorton et al reported that when weighted by degrees of freedom, the weighted mean r was .14, with a 95% confidence interval covering .10 to .19.

Forced-choice versus Free-response Results. Because forced-choice and free-response tests differ qualitatively, Honorton et al chose to examine their relationship to extraversion separately. They found that for free-response studies there was a significant correlation between extraversion and ESP scores, with mean $r = .20$ and $z = 4.46$. Further, this effect was homogeneous across both investigators and extraversion scales.

For forced-choice studies, there was a significant correlation between ESP and extraversion, but only for those studies that reported the ESP results to the subjects *before* measuring extraversion. Honorton et al speculated that the relationship was an artifact, in which extraversion scores were temporarily inflated as a result of positive feedback on ESP performance.

Confirmation with New Data. Following the extraversion/ESP meta-analysis, Honorton et al attempted to confirm the relationship using the autoganzfeld data base. Extraversion scores based on the Myers-Briggs Type Indicator were available for 221 of the 241 subjects who had participated in autoganzfeld studies.

The correlation between extraversion scores and ganzfeld rating scores was $r = .18$, with a 95% confidence interval from .05 to .30. This is consistent with the mean correlation of $r = .20$ for free-response experiments, determined from the meta-analysis. These correlations indicate that extraverted subjects can produce higher scores in free-response ESP tests.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

7. CONCLUSIONS

Parapsychologists often make a distinction between "proof-oriented research" and "process-oriented research." The former is typically conducted to test the hypothesis that psi abilities exist, while the latter is designed to answer questions about how psychic functioning works. Proof-oriented research has dominated the literature in parapsychology. Unfortunately, many of the studies used small samples and would thus be nonsignificant even if a moderate-sized effect exists.

The recent focus on meta-analysis in parapsychology has revealed that there are small but consistently nonzero effects across studies, experimenters, and laboratories. The size of the effects in forced-choice studies appear to be comparable to those reported in some medical studies that had been heralded as breakthroughs. (See Section 5, and Honorton and Ferrari, 1989, p. 301.) Free-response studies show effect sizes of far greater magnitude.

A promising direction for future process-oriented research is to examine the causes of individual differences in psychic functioning. The ESP/extraversion meta-analysis is a step in that direction.

In keeping with the idea of individual differences, Bayes and empirical Bayes methods would appear to make more sense than the classical inference methods commonly used, since they would allow individual abilities and beliefs to be modelled. Jeffreys (1990) reported a Bayesian analysis of some of the RNG experiments, and showed that conclusions were closely tied to prior beliefs even though hundreds of thousands of trials were available.

It may be that the nonzero effects observed in the meta-analyses can be explained by something other than ESP, such as shortcomings in our understanding of randomness and

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

independence. Nonetheless, there is an anomaly that needs an explanation. As I have argued elsewhere (Utts, 1987) research in parapsychology should receive more support from the scientific community. If ESP does not exist, there is little to be lost by erring in the direction of further research; which may in fact uncover other anomalies. If ESP does exist there is much to be lost by not doing process-oriented research, and much to be gained by discovering how to enhance and apply these abilities to important world problems.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

REFERENCES

Atkinson, R.L., Atkinson, R.C., Smith, E.E. and Bem, D.J. (1990). *Introduction to Psychology*, 10th Ed. Harcourt Brace Jovanovich, San Diego.

Beloff, J. (1985). Research strategies for dealing with unstable phenomena. In *The Repeatability Problem in Parapsychology* (B. Shapin and L. Coly. eds.) 1-21. Parapsychology Foundation, New York.

Blackmore, S.J. (1985). Unrepeatability: Parapsychology's only finding. In *The Repeatability Problem in Parapsychology* (B. Shapin and L. Coly. eds.) 183-206. Parapsychology Foundation, New York.

Burdick, D.S. and Kelly, E.F. (1977). Statistical methods in parapsychological research. In *Handbook of Parapsychology* (B.B. Wolman, ed.) 81-130. Van Nostrand Reinhold, New York.

Camp, B.H. (1937). (Statement in Notes Section.) *Journal of Parapsychology* **1** 305.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist* **45** 1304-1312.

Coover, J.E. (1917). *Experiments in Psychical Research at Leland Stanford Junior University*. Stanford University, Stanford, CA.

Dawes, R.M., Landman, J. and Williams, J. (1984). Reply to Kurosawa. *American Psychologist* **39** 74-75.

Diaconis, P. (1978). Statistical problems in ESP research. *Science* **201** 131-136.

Dommeyer, F.C. (1975). Psychical Research at Stanford University. *Journal of Parapsychology* **39** 173-205.

Druckman, D. and Swets, J.A., Eds. (1988). *Enhancing Human Performance: Issues, Theories, and Techniques*. National Academy Press, Washington, DC.

Edgeworth, F.Y. (1885). The calculus of probabilities applied to psychical research. *Proceedings of the Society for Psychical Research* **3** 190-199.

Edgeworth, F.Y. (1886). The calculus of probabilities applied to psychical research II. *Proceedings of the Society for Psychical Research* **4** 189-208.

Feller, W.K. (1968). *An Introduction to Probability Theory and Its Applications, Volume 1*, 3rd Ed. Wiley, New York.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Feller, W.K. (1940). Statistical aspects of ESP. *Journal of Parapsychology* **4** 271-297.

Fisher, R.A. (1924). A method of scoring coincidences in tests with playing cards. *Proceedings of the Society for Psychical Research* **34** 181-185.

Fisher, R.A. (1929). The statistical method in psychical research. *Proceedings of the Society for Psychical Research* **39** 189-192.

Gallup, G.H. Jr. and Newport, F. (1991). Belief in paranormal phenomena among adult Americans. *Skeptical Inquirer* **15** 137-146.

Gardner, M.J. and Altman, D.G. (1986). Confidence intervals rather than *p*-values: estimation rather than hypothesis testing. *British Medical Journal* **292** 746-750.

Gilmore, J.B. (1989). Randomness and the search for psi. *Journal of Parapsychology* **53** 309-340.

Gilmore, J.B. (1990). Anomalous significance in pararandom and psi-free domains. *Journal of Parapsychology* **54** 53-58.

Greeley, A. (1987). Mysticism goes mainstream. *American Health* **7** 47-49.

Greenhouse, J.B. and Greenhouse, S.W. (1988). An aspirin a day...? *Chance* **1** 24-31.

Greenwood, J.A. and Stuart, C.E. (1940). A review of Dr. Feller's critique. *Journal of Parapsychology* **4** 299-319.

Hansel, C.E.M. (1980). *ESP and Parapsychology: A Critical Re-evaluation*. Prometheus Books, Buffalo.

Harris, M.J. and Rosenthal, R. (1988a). *Interpersonal Expectancy Effects and Human Performance Research*. National Academy Press, Washington DC.

Harris, M.J. and Rosenthal, R. (1988b). *Postscript to Interpersonal Expectancy Effects and Human Performance Research*. National Academy Press, Washington DC.

Hedges, L.V. and Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Academic Press, Inc., Orlando, FL.

Honorton, C. (1977). Psi and internal attention states. In *Handbook of Parapsychology* (B.B. Wolman, ed.) 435-472. Van Nostrand Reinhold, New York.

Honorton, C. (1985a). How to evaluate and improve the replicability of parapsychological effects. In *The Repeatability Problem in Parapsychology* (B. Shapin and L. Coly. eds.)

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

238-255. Parapsychology Foundation, New York.

Honorton, C. (1985b). Meta-analysis of psi ganzfeld research: A response to Hyman. *Journal of Parapsychology* **49** 51-91.

Honorton, C., Berger, R.E., Vargoglis, M.P., Quant, M., Derr, P., Schechter, E.I. and Ferrari, D.C. (1990). Psi communication in the ganzfeld: experiments with an automated testing system and a comparison with a meta-analysis of earlier studies. *Journal of Parapsychology* **54** 99-139.

Honorton, C. and Ferrari, D.C. (1989). "Future telling": A meta-analysis of forced-choice precognition experiments, 1935-1987. *Journal of Parapsychology* **53** 281-308.

Honorton, C., Ferrari, D.C., and Bem, D.J. (1990). Extraversion and ESP Performance: A meta-analysis and a new confirmation. *Proceedings of the Annual Meeting of the Parapsychological Association.*

Hyman, R. (1985a). A critical overview of parapsychology. In *A Skeptic's Handbook of Parapsychology* (P. Kurtz, ed.) 1-96. Prometheus Books, Buffalo.

Hyman, R. (1985b). The ganzfeld psi experiment: A critical appraisal. *Journal of Parapsychology* **49** 3-49.

Hyman, R. and Honorton, C. (1986). Joint communiqué: The psi ganzfeld controversy. *Journal of Parapsychology* **50** 351-364.

Iversen, G.R, Longcor, W.H., Mosteller, F., Gilbert, J.P., and Youtz, C. (1971). Bias and runs in dice throwing and recording: A few million throws. *Psychometrika* **36** 1-19.

Jeffreys, W.H. (1990). Bayesian analysis of random event generator data. *Journal of Scientific Exploration* **4** 153-169.

Lindley, D.V. (1957). A statistical paradox. *Biometrika* **44** 187-192.

Mauskopf, S.H. and McVaugh, M. (1979). *The Elusive Science: Origins of Experimental Psychical Research.* The Johns Hopkins University Press, Baltimore.

McVaugh, M.R. and Mauskopf, S.H. (1976). J.B. Rhine's *Extrasensory Perception* and its background in psychical research. *Isis* **67** 161-189.

Neuliep, J.W. (Ed.) (1990). Handbook of replication research in the behavioral and social sciences. *Journal of Social Behavior and Personality* (Special Issue) **5**(4).

Office of Technology Assessment (1989). Report of a workshop on experimental

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

parapsychology. *Journal of the American Society for Psychical Research* **83** 317-339.

Palmer, J. (1989). A reply to Gilmore. *Journal of Parapsychology* **53** 341-344.

Palmer, J. (1990). Reply to Gilmore: Round two. *Journal of Parapsychology* **54** 59-61.

Palmer, J.A., Honorton, C. and Utts, J. (1989). Reply to the National Research Council study on parapsychology. *Journal of the American Society for Psychical Research* **83** 31-49.

Radin, D.I. and Ferrari, D.C. (1991). Effects of consciousness on the fall of dice: A meta-analysis. *Journal of Scientific Exploration* **5** (to appear).

Radin, D.I. and Nelson, R.D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics* **19** 1499-1514.

Rao, K.R. (1985). Replication in conventional and controversial sciences. In *The Repeatability Problem in Parapsychology* (B. Shapin and L. Coly. eds.) 22-41. Parapsychology Foundation, New York.

Rhine, J.B. (1934). *Extrasensory Perception*. Boston Society for Psychical Research, Boston. (Reprinted by Branden Press in 1964).

Rhine, J.B. (1977). History of experimental studies. In *Handbook of Parapsychology* (B.B. Wolman, ed.) 25-47. Van Nostrand Reinhold, New York.

Richet, C. (1884). La suggestion mentale et le calcul des probabilites. *Revue Philosophique* **18** 608-674.

Rosenthal, R. (1984). *Meta-Analytic Procedures for Social Research.* Sage, Beverly Hills.

Rosenthal, R. (1986). Meta-analytic procedures and the nature of replication: The ganzfeld debate. *Journal of Parapsychology* **50** 315-336.

Rosenthal, R. (1990a). How are we doing in soft psychology? *American Psychologist* **45** 775-777.

Rosenthal, R. (1990b). Replication in behavioral research. *Journal of Social Behavior and Personality* **5** 1-30.

Saunders, D.R. (1985). On Hyman's factor analysis. *Journal of Parapsychology* **49** 86-88.

Shapin, B. and Coly, L. (Eds.) (1985). *The Repeatability Problem in Parapsychology.* Parapsychology Foundation, New York.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Spencer-Brown, G. (1957). *Probability and Scientific Inference.* Longmans Green, London and New York.

Stuart, C.E. and Greenwood, J.A. (1937). A review of criticisms of the mathematical evaluation of ESP data. *Journal of Parapsychology* 1 295-304.

Tversky, A. and Kahneman, D. (1982). Belief in the law of small numbers. In D. Kahneman, P. Slovic and A. Tversky (Eds), *Judgment under uncertainty: Heuristics and biases.* Cambridge University Press, Cambridge.

Utts, J. (1986). The ganzfeld debate: A statistician's perspective. *Journal of Parapsychology* **50** 395-402.

Utts, J. (1987). Psi, statistics, and society. *Behavioral and Brain Sciences* **10** 615-616.

Utts, J. (1988). Successful replication versus statistical significance. *Journal of Parapsychology* **52** 305-320.

Utts, J. (1989a). Randomness and randomization tests: A reply to Gilmore. *Journal of Parapsychology* **53** 345-351.

Utts, J. (1989b). Analyzing free-response data - a progress report. To appear in *Psi Research Methodology: A Re-examination* (L. Coly, ed.). Parapsychology Foundation, New York.

Wilks, S.S. (1965). *N.Y. Statistician* **16** (nos. 6 and 7).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

III MAIN-STREAM PUBLICATIONS

One measure of the acceptance of anomalous mental phenomena as a valid area for investigation is the degree to which research papers appear in the main-stream scientific literature. The reports in this section have been selected because they are a representative sample of such papers.

The number that appears in the upper right–hand corner of the first page for each publication is keyed to the following descriptions:

9. Targ, R. and Puthoff, H. E., "Information Transmission Under Conditions of Sensory Shielding," *Nature,* Vol. 252, pp. 602-607, (October, 1974). Targ and Puthoff describe a series of experiments with selected individuals, including Mr. Uri Geller, and introduce an anomalous cognition technique called *remote viewing.* The paper also includes a pilot experiment to investigate the effects of anomalous cognition on the alpha rhythms in the brain.
10. Puthoff, H. E. and Targ, R., "A Perceptual Channel for Information Transfer over Kilometer Distances: Historical Perspective and Recent Research," *Proceedings of the IEEE,* Vol. 64, No. 3, pp. 329-354, (March, 1976). Puthoff and Targ provide a historical review of the pertinent literature and describe over 50 remote viewing (i.e., anomalous cognition) trials. The paper also includes representative examples of remote viewing.
11. Jahn, R. G., "The Persistent Paradox of Psychic Phenomena: An Engineering Perspective," Invited Paper, *Proceedings of the IEEE,* Vol. 70, No. 2, pp. 136-170, (February, 1982). Jahn describes a replication of remote viewing and extends the distance to over 10,000 kilometers. In addition to an independent overview of parapsychology, Jahn also includes descriptions of a number of anomalous perturbation experiments.
12. Child, I. L., "Psychology and Anomalous Observations: The Question of ESP in Dreams," *American Psychologist,* Vol. 40, No. 11, pp. 1219-1230, (November, 1985). Professor Child, the then Chairman of the Psychology Department at Yale University, provides a critical review of the anomalous cognition dream studies conducted at Maimonides Medical Center in the early 1970's. Professor Child warns the general psychological research community not to dismiss the body of research and suggests that it should be of wide interest to them.
13. Atkinson, R. L., Atkinson, R. C., Smith, E. E., and Bem, D. J., *Introduction to Psychology, 10th Edition,* pp. 234-243, Harcourt Brace Jovanovich, New York, (1990). Professor Bem included anomalous cognition in a chapter on consciousness and its altered states in a widely-used introductory text in psychology. Bem provides definitions of terms, a review of the experimental evidence for anomalous cognition, an analysis of the debate over the evidence, and a review of the anecdotal evidence.
14. Walker, E. H., May, E. C., Spottiswoode, S. J. P., and Piantanida, T., "Testing Schrödinger's Paradox with a Michelson Interferometer," *Physics B,* Vol. 151, pp. 339-348, (1988). While not directly related to anomalous mental phenomena, this paper describes an experimental test to determine if consciousness is a *necessary* ingredient for determining physical reality. The authors conclude that is it not, and thus, this result has implications for anomalous perturbation research.
15. Hyman, R., "Parapsychological Research: A Tutorial Review and Critical Appraisal," Invited Paper, *Proceedings of the IEEE,* Vol. 74, No. 6, pp. 823-849, (June, 1986). Dr. Hyman is a Professor of Psychology at the University of Oregon in Eugene and has been a long-time critic of and commentator on the field of parapsychology. Hyman reviews the historical experiments and provides a critical analysis of the current research.

Approved For Release 20

Vol. 251 No. 5476 October 18, 1974

Published weekly by
Macmillan Journals Ltd

London
4 Little Essex Street, WC2R 3LF
Telephone: (01) 836 6633 Telex: 262024
Telegrams: Phusis London WC2R 3LF

Washington
711 National Press Building, DC 20045
Telephone: (202) 737 2355
Telex: 64280

Editor
David Davies

Deputy Editor
Roger Woodham

Editorial staff
Gillian Boucher, Colin Norman*, John Gribbin, Sally Owen, John Hall, Allan Piper, Eleanor Lawrence, Miranda Robertson, *Mary Lindley, Fiona Selkirk, Peter Milford, Robert Vickers, Peter Newmark, Mary Wade*, John Wilson
*Washington office

Publishing Director
Jenny Hughes

Display advertisement enquiries to:
London Office

Classified advertisement enquiries to:
T. G. Scott and Son Ltd,
1 Clement's Inn,
London WC2A 2ED
Telephone: (01) 242 6264 and
(01) 405 4743
Telegrams: Textualist London
WC2A 2ED

Subscription enquiries to:
Macmillan Journals Ltd, Brunel Road,
Basingstoke, Hants, RG21 2XS
Telephone: Basingstoke 29242

Publication address in the United States
The Wm Byrd Press Inc.,
2901 Byrdhill Road,
Richmond, Virginia 23228

Second Class Postage for the USA
paid at Richmond, Virginia

US Postmaster, please send form 3579
to Nature, 711 National Press Building,
Washington DC 20045

Price
£22 per year—excepting USA
and Canada (£28 per year)
Registered as a newspaper at the
British Post Office

Copyright © Macmillan Journals Ltd,
October 18, 1974

Cover Picture

A hundred years ago *Nature* was reviewing E. J. Marey's *Animal Mechanism* (page 518, October 29, 1874). These cumbersome mechanisms were soon to be replaced by Muybridge's zoopraxiscope camera. On page 567 we look at a Muybridge sequence and—a century later—what happens when the light is switched on.

9

nature

Volume 252 October 18, 1974

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Information transmission under conditions of sensory shielding

WE present results of experiments suggesting the existence of one or more perceptual modalities through which individuals obtain information about their environment, although this information is not presented to any known sense. The literature[1-3] and our observations lead us to conclude that such abilities can be studied under laboratory conditions.

We have investigated the ability of certain people to describe graphical material or remote scenes shielded against ordinary perception. In addition, we performed pilot studies to determine if electroencephalographic (EEG) recordings might indicate perception of remote happenings even in the absence of correct overt responses.

We concentrated on what we consider to be our primary responsibility—to resolve under conditions as unambiguous as possible the basic issue of whether a certain class of paranormal perception phenomena exists. So we conducted our experiments with sufficient control, utilising visual, acoustic and electrical shielding, to ensure that all conventional paths of sensory input were blocked. At all times we took measures to prevent sensory leakage and to prevent deception, whether intentional or unintentional.

Our goal is not just to catalogue interesting events, but to uncover patterns of cause-effect relationships that lend themselves to analysis and hypothesis in the forms with which we are familiar in scientific study. The results presented here constitute a first step towards that goal; we have established under known conditions a data base from which departures as a function of physical and psychological variables can be studied in future work.

REMOTE PERCEPTION OF GRAPHIC MATERIAL

First, we conducted experiments with Mr Uri Geller in which we examined his ability, while located in an electrically shielded room, to reproduce target pictures drawn by experimenters located at remote locations. Second, we conducted double-blind experiments with Mr Pat Price, in which we measured his ability to describe remote outdoor scenes many miles from his physical location. Finally, we conducted pre-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

liminary tests using EEGs, in which subjects were asked to perceive whether a remote light was flashing, and to determine whether a subject could perceive the presence of the light, even if only at a noncognitive level of awareness.

In preliminary testing Geller apparently demonstrated an ability to reproduce simple pictures (line drawings) which had been drawn and placed in opaque sealed envelopes which he was not permitted to handle. But since each of the targets was known to at least one experimenter in the room with Geller, it was not possible on the basis of the preliminary testing to discriminate between Geller's direct perception of envelope contents and perception through some mechanism involving the experimenters, whether paranormal or subliminal.

So we examined the phenomenon under conditions designed to eliminate all conventional information channels, overt or subliminal. Geller was separated from both the target material and anyone knowledgeable of the material, as in the experiments of ref. 4.

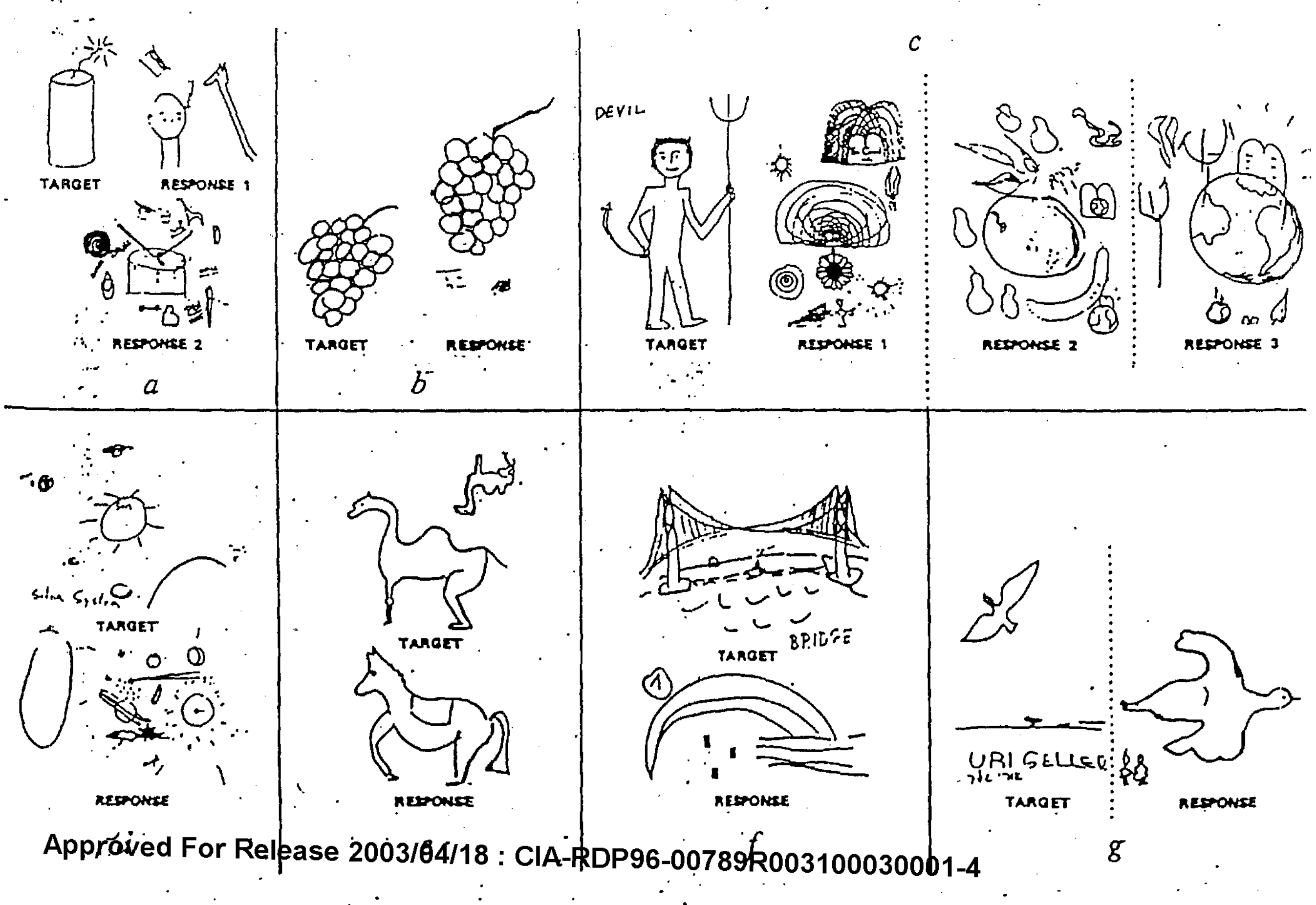
In the first part of the study a series of 13 separate drawing experiments were carried out over 7 days. No experiments are deleted from the results presented here.

At the beginning of the experiment either Geller or the experimenters entered a shielded room so that from that time forward Geller was at all times visually, acoustically and electrically shielded from personnel and material at the target location. Only following Geller's isolation from the experimenters was a target chosen and drawn, a procedure designed to eliminate pre-experiment cueing. Furthermore, to eliminate the possibility of pre-experiment target forcing, Geller was kept ignorant as to the identity of the person selecting the target and as to the method of target selection. This was accomplished by the use of three different techniques: (1) pseudo-random technique of opening a dictionary arbitrarily and choosing the first word that could be drawn (Experiments 1–4); (2) targets, blind to experimenters and subject, prepared independently by SRI scientists outside the experimental group (followin Geller's isolation) and provided to the experimenters durin the course of the experiment (Experiments 5–7, 11–13); and (3 arbitrary selection from a target pool decided upon in advanc of daily experimentation and designed to provide data concern ing information content for use in testing specific hypothese (Experiments 8–10). Geller's task was to reproduce with pe on paper the line drawing generated at the target location Following a period of effort ranging from a few minutes t half an hour, Geller either passed (when he did not feel con fident) or indicated he was ready to submit a drawing to th experimenters, in which case the drawing was collected befor Geller was permitted to see the target.

To prevent sensory cueing of the target information, Experiment 1 through 10 were carried out using a shielded room in SRI's facilit for EEG research. The acoustic and visual isolation is provide by a double-walled steel room, locked by means of an inner an outer door, each of which is secured with a refrigerator-type lockin mechanism. Following target selection when Geller was insid the room, a one-way audio monitor, operating only from the insid to the outside, was activated to monitor Geller during his efforts The target picture was never discussed by the experimenters after th picture was drawn and brought near the shielded room. In ou detailed examination of the shielded room and the protocol used i these experiments, no sensory leakage has been found.

The conditions and results for the 10 experiments carried out in th shielded room are displayed in Table 1 and Fig. 1. All experimen except 4 and 5, were conducted with Geller inside the shielded room In Experiments 4 and 5, the procedure was reversed. For thos experiments in which Geller was inside the shielded room, the targ location was in an adjacent room at a distance of about 4 m, excep for Experiments 3 and 8, in which the target locations were, respe tively, an office at a distance of 475 m and a room at a distance o about 7 m.

A response was obtained in all experiments except Number 5–7. In Experiment 5, the person-to-person link was eliminate by arranging for a scientist outside the usual experimenta group to draw a picture, lock it in the shielded room befor Geller's arrival at SRI, and leave the area. Geller was then le



Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Fig. 1 Target pictures and responses drawn by Uri Geller under shielded conditions.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 1 Remote perception of graphic material

Experiment	Date (month, day, year)	Geller Location	Target Location	Target	Figure
1	8/4/73	Shielded room 1*	Adjacent room (4.1 m)†	Firecracker	1*a*
2	8/4/73	Shielded room 1	Adjacent room (4.1 m)	Grapes	1*b*
3	8/5/73	Shielded room 1	Office (475 m)	Devil	1*c*
4	8/5/73	Room adjacent to shielded room 1	Shielded room 1 (3.2 m)	Solar system	1*d*
5	8/6/73	Room adjacent to shielded room 1	Shielded room 1 (3.2 m)	Rabbit	No drawing
6	8/7/73	Shielded room 1	Adjacent room (4.1 m)	Tree	No drawing
7	8/7/73	Shielded room 1	Adjacent room (4.1 m)	Envelope	No drawing
8	8/8/73	Shielded room 1	Remote room (6.75 m)	Camel	1*e*
9	8/8/73	Shielded room 1	Adjacent room (4.1 m)	Bridge	1*f*
10	8/8/73	Shielded room 1	Adjacent room (4.1 m)	Seagull	1*g*
11	8/9/73	Shielded room 2‡	Computer (54 m)	Kite (computer CRT)	2*a*
12	8/10/73	Shielded room 2	Computer (54 m)	Church (computer memory)	2*b*
13	8/10/73	Shielded room 2	Computer (54 m)	Arrow through heart (computer CRT, zero intensity)	2*c*

*EEG Facility shielded room (see text).
†Perceiver–target distances measured in metres.
‡SRI Radio Systems Laboratory shielded room (see text).

by the experimenters to the shielded room and asked to draw the picture located inside the room. He said that he got no clear impression and therefore did not submit a drawing. The elimination of the person-to-person link was examined further in the second series of experiments with this subject.

Experiments 6 and 7 were carried out while we attempted to record Geller's EEG during his efforts to perceive the target pictures. The target pictures were, respectively, a tree and an envelope. He found it difficult to hold adequately still for good EEG records, said that he experienced difficulty in getting impressions of the targets and again submitted no drawings.

Experiments 11 through 13 were carried out in SRI's Engineering Building, to make use of the computer facilities available there. For these experimenters, Geller was secured in a double-walled, copper-screen Faraday cage 54 m down the hall and around the corner from the computer room. The Faraday cage provides 120 dB attenuation for plane wave radio frequency radiation over a range of 15 kHz to 1 GHz. For magnetic fields the attenuation is 68 dB at 15 kHz and decreases to 3 dB at 60 Hz. Following Geller's isolation, the targets for these experiments were chosen by computer laboratory personnel not otherwise associated with either the experiment or Geller, and the experimenters and subject were kept blind as to the contents of the target pool.

For Experiment 11, a picture of a kite was drawn on the face of a cathode ray tube display screen, driven by the computer's graphics program. For Experiment 12, a picture of a church was drawn and stored in the memory of the computer. In Experiment 13, the target drawing, an arrow through a heart (Fig. 2*c*), was drawn on the face of the cathode ray tube and then the display intensity was turned off so that no picture was visible.

To obtain an independent evaluation of the correlation between target and response data, the experimenters submitted the data for judging on a 'blind' basis by two SRI scientists who were not otherwise associated with the research. For the 10 cases in which Geller provided a response, the judges were asked to match the response data with the corresponding target data (without replacement). In those cases in which Geller made more than one drawing as his response to the target, all the drawings were combined as a set for judging. The two judges each matched the target data to the response data with no error. For either judge such a correspondence has an *a priori* probability, under the null hypothesis of no information channel, of $P = (10!)^{-1} = 3 \times 10^{-7}$.

A second series of experiments was carried out to determine whether direct perception of envelope contents was possible without some person knowing of the target picture.

One hundred target pictures of everyday objects were drawn by an SRI artist and sealed by other SRI personnel in double envelopes containing black cardboard. The hundred targets were divided randomly into groups of 20 for use in each of the three days' experiments.

On each of the three days of these experiments, Geller passed. That is, he declined to associate any envelope with a drawing that he made, expressing dissatisfaction with the existence of such a large target pool. On each day he made approximately 12 recognisable drawings, which he felt were associated with the entire target pool of 100. On each of the three days, two of his drawings could reasonably be associated with two of the 20 daily targets. On the third day, two of his drawings were very close replications of two of that day's target pictures. The drawings resulting from this experiment do not depart significantly from what would be expected by chance.

In a simpler experiment Geller was successful in obtaining information under conditions in which no persons were knowledgeable of the target. A double-blind experiment was performed in which a single 3/4 inch die was placed in a 3 × 4 × 5 inch steel box. The box was then vigorously shaken by one of the experimenters and placed on the table, a technique found in control runs to produce a distribution of die faces differing nonsignificantly from chance. The orientation of the die within the box was unknown to the experimenters at that time. Geller would then write down which die face was uppermost. The target pool was known, but the targets were individually prepared in a manner blind to all persons involved in the experiment. This experiment was performed ten times, with Geller passing twice and giving a response eight times. In the eight times in which he gave a response, he was correct each time. The distribution of responses consisted of three 2s, one 4, two 5s, and two 6s. The probability of this occurring by chance is approximately one in 10^6.

In certain situations significant information transmission can take place under shielded conditions. Factors which appear to be important and therefore candidates for future investigation include whether the subject knows the set of targets in the target pool, the actual number of targets in the target pool at any given time, and whether the target is known by any of the experimenters.

It has been widely reported that Geller has demonstrated the ability to bend metal by paranormal means. Although metal bending by Geller has been observed in our laboratory, we have not been able to combine such observations with adequately controlled experiments to obtain data sufficient to support the paranormal hypothesis.

REMOTE VIEWING OF NATURAL TARGETS

A study by Osis[5] led us to determine whether a subject could describe randomly chosen geographical sites located several miles from the subject's position and demarcated by some

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

appropriate means (remote viewing). This experiment carried out with Price, a former California police commissioner and city councilman, consisted of a series of double-blind, demonstration-of-ability tests involving local targets in the San Francisco Bay area which could be documented by several independent judges. We planned the experiment considering that natural geographical places or man-made sites that have existed for a long time are more potent targets for paranormal perception experiments than are artificial targets prepared in the laboratory. This is based on subject opinions that the use of artificial targets involves a 'trivialisation of the ability' as compared with natural pre-existing targets.

In each of nine experiments involving Price as subject and SRI experimenters as a target demarcation team, a remote location was chosen in a double-blind protocol. Price, who remained at SRI, was asked to describe this remote location, as well as whatever activities might be going on there.

Several descriptions yielded significantly correct data pertaining to and descriptive of the target location.

In the experiments a set of twelve target locations clearly differentiated from each other and within 30 min driving time from SRI had been chosen from a target-rich environment (more than 100 targets of the type used in the experimental series) prior to the experimental series by an individual in SRI management, the director of the Information Science and Engineering Division, not otherwise associated with the experiment. Both the experimenters and the subject were kept blind as to the contents of the target pool, which were used without replacement.

An experimenter was closeted with Price at SRI to wait 30 min to begin the narrative description of the remote location. The SRI locations from which the subject viewed the remote locations consisted of an outdoor park (Experiments 1, 2), the double-walled copper-screen Faraday cage discussed earlier (Experiments 3, 4, and 6–9); and an office (Experiment 5). A second experimenter would then obtain a target location from the Division Director from a set of travelling orders previously prepared and randomised by the Director and kept under his control. The target demarcation team (two to four SRI experimenters) then proceeded directly to the target by automobile without communicating with the subject or experimenter remaining behind. Since the experimenter remaining with the subject at SRI was in ignorance both as to the particular target and as to the target pool, he was free to question Price to clarify his descriptions. The demarcation team then remained at the target site for 30 min after the 30 min allotted for travel. During the observation period, the remote-viewing subject would describe his impressions of the target site into a tape recorder. A comparison was then made when the demarcation team returned.

Price's ability to describe correctly buildings, docks, roads, gardens and so on, including structural materials, colour, ambience and activity, sometimes in great detail, indicated the functioning of a remote perceptual ability. But the descriptions contained inaccuracies as well as correct statements. To obtain a numerical evaluation of the accuracy of the remote viewing experiment, the experimental results were subjected to independent judging on a blind basis by five SRI scientists who were

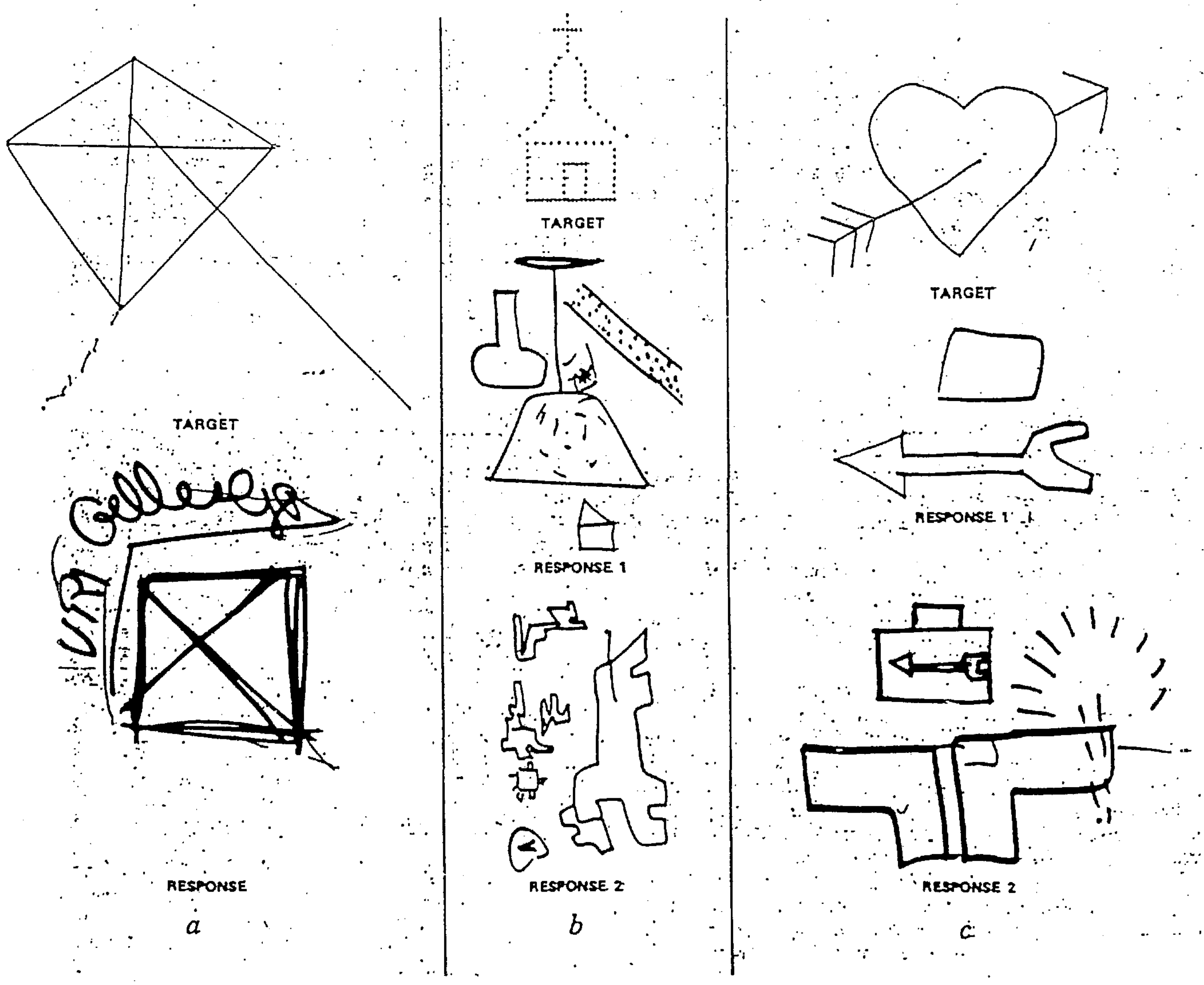


Fig. 2. Computer drawings and responses drawn by Uri Geller. *a*, Computer drawing stored on video display; *b*, computer drawing stored in computer memory only; *c*, computer drawing stored on video display with zero intensity.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 2 Distribution of correct selections by judges A, B, C, D, and E in remote viewing experiments

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Descriptions chosen by judges		Places visited by judges 1	2	3	4	5	6	7	8	9
Hoover Tower	1	ABCDE				D				
Baylands Nature Preserve	2		ABC	E				D		D
Radio Telescope	3			ACD		BE				
Redwood City Marina	4		CD		ABDE		E			
Bridge Toll Plaza	5						ABD		DCE	
Drive-In Theatre	6			B		A	C			E
Arts and Crafts Garden Plaza	7							ABCE		
Church	8				C				AB	
Rinconada Park	9		CE							A[illegible]

Of the 45 selections (5 judges, 9 choices), 24 were correct. Bold type indicates the description chosen most often for each place visited. C[illegible] choices lie on the main diagonal. The number of correct matches by Judges A through E is 7, 6, 5, 3, and 3, respectively. The expected nu[illegible] of correct matches from the five judges was five; in the experiment 24 such matches were obtained. The *a priori* probability of such an occur[illegible] by chance, conservatively assuming assignment without replacement on the part of the judges, is $P = 8 \cdot 10^{-10}$.

not otherwise associated with the research. The judges were asked to match the nine locations, which they independently visited, against the typed manuscripts of the tape-recorded narratives of the remote viewer. The transcripts were unlabelled and presented in random order. The judges were asked to find a narrative which they would consider the best match for each of the places they visited. A given narrative could be assigned to more than one target location. A correct match requires that the transcript of a given date be associated with the target of that date. Table 2 shows the distribution of the judges' choices.

Among all possible analyses, the most conservative is a permutation analysis of the plurality vote of the judges' selections assuming assignment without replacement, an approach independent of the number of judges. By plurality vote, six of the nine descriptions and locations were correctly matched. Under the null hypothesis (no remote viewing and a random selection of descriptions without replacement), this outcome has an *a priori* probability of $P = 5.6 \times 10^{-4}$, since, among all possible permutations of the integers one through nine, the probability of six or more being in their natural position in the list has that value. Therefore, although Price's descriptions contain inaccuracies, the descriptions are sufficiently accurate to permit the judges to differentiate among the various targets to the degree indicated.

EEG EXPERIMENTS

An experiment was undertaken to determine whether a physiological measure such as EEG activity could be used as an indicator of information transmission between an isolated subject and a remote stimulus. We hypothesised that perception could be indicated by such a measure even in the absence of verbal or other overt indicators[6,7].

It was assumed that the application of remote stimuli would result in responses similar to those obtained under conditions of direct stimulation. For example, when normal subjects are stimulated with a flashing light, their EEG typically shows a decrease in the amplitude of the resting rhythm and a driving of the brain waves at the frequency of the flashes[8]. We hypothesised that if we stimulated one subject in this manner (a sender), the EEG of another subject in a remote room with no flash present (a receiver), might show changes in alpha (9–11 Hz) activity, and possibly EEG driving similar to that of the sender.

We informed our subject that at certain times a light was to be flashed in a sender's eyes in a distant room, and if the subject perceived that event, consciously or unconsciously, it might be evident from changes in his EEG output. The receiver was seated in the visually opaque, acoustically and electrically shielded double-walled steel room previously described. The sender was seated in a room about 7 m from the receiver.

To find subjects who were responsive to such a remote stimulus, we initially worked with four female and two male volunteer subjects, all of whom believed that success in the experimental situation might be possible. These were designated 'receivers'. The senders were either other subjects or [illegible] experimenters. We decided beforehand to run one or tw[illegible] sessions of 36 trials each with each subject in this sele[illegible] procedure, and to do a more extensive study with any su[illegible] whose results were positive.

A Grass PS-2 photostimulator placed about 1 m in front of t[illegible] sender was used to present flash trains of 10 s duration. The receive[illegible] EEG activity from the occipital region (O_z), referenced to [illegible] mastoids, was amplified with a Grass 5P-1 preamplifier and asso[illegible] driver amplifier with a bandpass of 1–120 Hz. The EEG data we[illegible] recorded on magnetic tape with an Ampex SP 300 recorder.

On each trial, a tone burst of fixed frequency was presented to bo[illegible] sender and receiver and was followed in one second by either [illegible] train of flashes or a null flash interval presented to the sender. T[illegible] six such trials were given in an experimental session, consisting o[illegible] null trials—no flashes following the tone—12 trials of flashes at 6 f.p[illegible] and 12 trials of flashes at 16 f.p.s., all randomly intermixed, dete[illegible] mined by entries from a table of random numbers. Each of the [illegible] generated an 11-s EEG epoch. The last 4 s of the epoch was se[illegible] for analysis to minimise the desynchronising action of the warni[illegible] cue. This 4-s segment was subjected to Fourier analysis on a LINC computer.

Spectrum analyses gave no evidence of EEG driving in any re[illegible] although in control runs the receivers did exhibit driving [illegible] physically stimulated with the flashes. But of the six subjects stud[illegible] initially, one subject (H. H.) showed a consistent alpha blocking effe[illegible] We therefore undertook further study with this subject.

Data from seven sets of 36 trials each were collected fro[illegible] subject on three separate days. This comprises all the data col[illegible] to date with this subject under the test conditions described ab[illegible] The alpha band was identified from average spectra, then scores [illegible] average power and peak power were obtained from individual tria[illegible] and subjected to statistical analysis.

Of our six subjects, H. H. had by far the most monochro[illegible] EEG spectrum. Figure 3 shows an overlay of the three averag[illegible] spectra from one of this subject's 36-trial runs, displayi[illegible] changes in her alpha activity for the three stimulus condi[illegible]

Mean values for the average power and peak power fo[illegible]

Table 3 EEG data for H.H. showing average power and peak [illegible] in the 9–11 Hz band, as a function of flash frequency and [illegible]

Flash Frequency	0	6	16	0	6	16
Sender	Average Power			Peak Power		
J.L.	94.8	84.1	76.8	357.7	329.2	289.6
R.T.	41.3	45.5	37.0	160.7	161.0	125.0
No sender (subject informed)	25.1	35.7	28.2	87.5	95.7	81.7
J.L.	54.2	55.3	44.8	191.4	170.5	149.3
J.L.	56.8	50.9	32.8	240.6	178.0	104.6
R.T.	39.8	24.9	30.3	145.2	74.2	122.1
No sender (subject not informed)	86.0	53.0	52.1	318.1	180.6	202.3
Averages	56.8	49.9	43.1	214.5	169.8	153.5
		−12%	−24% ($P < 0.04$)		−21%	−28% ($P <$ 0.0[illegible]

Each entry is an average over 12 trials

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

of the seven experimental sets are given in Table 3. The power measures were less in the 16 f.p.s. case than in the 0 f.p.s. in all seven peak power measures and in six out of seven average power measures. Note also the reduced effect in the case in which the subject was informed that no sender was present (Run 3). It seems that overall alpha production was reduced for this run in conjunction with the subject's expressed apprehension about conducting the experiment without a sender. This is in contrast to the case (Run 7) in which the subject was not informed.

Siegel's two-tailed t approximation to the nonparametric randomisation test[9] was applied to the data from all sets, which included two sessions in which the sender was removed. Average power on trials associated with the occurrence of 16 f.p.s. was significantly less than when there were no flashes ($t = 2.09$, d.f. = 118, $P<0.04$). The second measure, peak power, was also significantly less in the 16 f.p.s. conditions than in the null condition ($t = 2.16$, d.f. = 118, $P<0.03$). The average response in the 6 f.p.s. condition was in the same direction as that associated with 16 f.p.s., but the effect was not statistically significant.

Spectrum analyses of control recordings made from saline with a 12 kΩ resistance in place of the subject with and without the addition of a 10 Hz, 50 μV test signal applied to the saline solution, revealed no indications of flash frequencies, nor perturbations of the 10 Hz signal. These controls suggest that the results were not due to system artefacts. Further tests also gave no evidence of radio frequency energy associated with the stimulus.

Subjects were asked to indicate their conscious assessment for each trial as to which stimulus was generated. They made their guesses known to the experimenter via one-way telegraphic communication. An analysis of these guesses has shown them to be at chance, indicating the absence of any supraliminal cueing, so arousal as evidenced by significant alpha blocking occurred only at the noncognitive level of awareness.

We hypothesise that the protocol described here may prove to be useful as a screening procedure for latent remote perceptual ability in the general population.

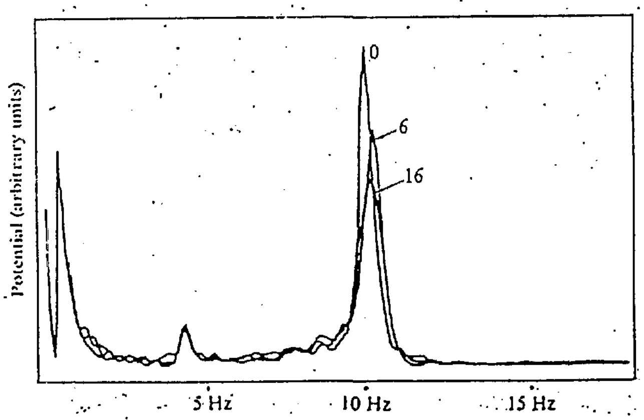


Fig. 3 Occipital EEG spectra, 0–20 Hz, for one subject (H. H.) acting as receiver, showing amplitude changes in the 9–11 Hz band as a function of strobe frequency. Three cases: 0, 6, and 16 f.p.s. (12 trial averages).

CONCLUSION

From these experiments we conclude that:

- A channel exists whereby information about a remote location can be obtained by means of an as yet unidentified perceptual modality.
- As with all biological systems, the information channel appears to be imperfect, containing noise along with the signal.
- While a quantitative signal-to-noise ratio in the information-theoretical sense cannot as yet be determined, the results of our experiments indicate that the functioning is at the level of useful information transfer.

It may be that remote perceptual ability is widely distributed in the general population, but because the perception is generally below an individual's level of awareness, it is repressed or not noticed. For example, two of our subjects (H. H. and P. P.) had not considered themselves to have unusual perceptual ability before their participation in these experiments.

Our observation of the phenomena leads us to conclude that experiments in the area of so-called paranormal phenomena can be scientifically conducted, and it is our hope that other laboratories will initiate additional research to attempt to replicate these findings.

This research was sponsored by The Foundation for Parasensory Investigation, New York City. We thank Mrs Judith Skutch, Dr Edgar D. Mitchell of the Institute of Noetic Sciences—as well as our SRI associates, Mr Bonnar Cox, Mr Earle Jones and Dr Dean Brown—for support and encouragement. Constructive suggestions by Mrs Jean Mayo, Dr Charles Tart, University of California, and Dr Robert Ornstein and Dr David Galin of the Langley Porter Neuropsychiatric Institute are acknowledged.

RUSSELL TARG
HAROLD PUTHOFF

Electronics and Bioengineering Laboratory,
Stanford Research Institute,
Menlo Park, California 94025

Received March 11; revised July 8, 1974.

1. Pratt, J., Rhine, J. B., Stuart, C., and Greenwood, J., *Extra-Sensory Perception after Sixty Years* (Henry Holt, New York, 1940).
2. Soal, S., and Bateman, F., *Modern Experiments in Telepathy* (Faber and Faber, London, 1954).
3. Vasilliev, L. L., *Experiments in Mental Suggestion* (ISMI Publications, Hampshire, England, 1963).
4. Musso, J. R., and Granero, M., *J. Parapsychology*, 37, 13–3 (1973).
5. Osis, K., *ASPR Newsletter*, No. 14 (1972).
6. Tart, C. T., *Physiological Correlates of Psi Cognition*, *Int. J. Parapsychology*, V, No. 4 (1963).
7. Dean, E. D., *Int. J. Neuropsychiatry*, 2 (1966).
8. Hill, D., and Parr, G., *Electroencephalography: A Symposium on its Various Aspects* (Macmillan, New York, 1963).
9. Siegel, S., *Nonparametric Statistics for the Behavioral Sciences*, 152–156 (McGraw-Hill, New York, 1956).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

PROCEEDINGS OF THE IEEE, VOL. 64, NO. 3, MARCH 1976
Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

A Perceptual Channel for Information Transfer over Kilometer Distances: Historical Perspective and Recent Research

HAROLD E. PUTHOFF, MEMBER, IEEE, AND RUSSELL TARG, SENIOR MEMBER, IEEE

Abstract—For more than 100 years, scientists have attempted to determine the truth or falsity of claims for the existence of a perceptual channel whereby certain individuals are able to perceive and describe remote data not presented to any known sense. This paper presents an outline of the history of scientific inquiry into such so-called paranormal perception and surveys the current state of the art in parapsychological research in the United States and abroad. The nature of this perceptual channel is examined in a series of experiments carried out in the Electronics and Bioengineering Laboratory of Stanford Research Institute. The perceptual modality most extensively investigated is the ability of both experienced subjects and inexperienced volunteers to view, by innate mental processes, remote geographical or technical targets including buildings, roads, and laboratory apparatus. The accumulated data indicate that the phenomenon is not a sensitive function of distance, and Faraday cage shielding does not in any apparent way degrade the quality and accuracy of perception. On the basis of this research, some areas of physics are suggested from which a description or explanation of the phenomenon could be forthcoming.

I. Introduction

"IT IS THE PROVINCE of natural science to investigate nature, impartially and without prejudice" [1]. Nowhere in scientific inquiry has this dictum met as great a challenge as in the area of so-called extrasensory perception (ESP), the detection of remote stimuli not mediated by the usual sensory processes. Such phenomena, although under scientific consideration for over a century, have historically been fraught with unreliability and controversy, and validation of the phenomena by accepted scientific methodology has been slow in coming. Even so, a recent survey conducted by the British publication *New Scientist* revealed that 67 percent of nearly 1500 responding readers (the majority of whom are working scientists and technologists) considered ESP to be an established fact or a likely possibility, and 88 percent held the investigation of ESP to be a legitimate scientific undertaking [2].

A review of the literature reveals that although experiments by reputable researchers yielding positive results were begun over a century ago (e.g., Sir William Crookes' study of D. D. Home, 1860's) [3], many consider the study of these phenomena as only recently emerging from the realm of quasi-science. One reason for this is that, despite experimental results, no satisfactory theoretical construct had been advanced to correlate data or to predict new experimental outcomes. Consequently, the area in question remained for a long time in the recipe stage reminiscent of electrodynamics before the unification brought about by the work of Ampere, Faraday, and Maxwell. Since the early work, however, we have seen the development of information theory, quantum theory, and neurophysiological research, and these disciplines provide powerful conceptual tools that appear to bear directly on the issue. In fact, several physicists (Section V) are now of the opinion that these phenomena are not at all inconsistent with the framework of modern physics: the often-held view that observations of this type are *a priori* incompatible with known laws is erroneous in that such a concept is based on the naive realism prevalent before the development of quantum theory. In the emerging view, it is accepted that research in this area can be conducted so as to uncover not just a catalog of interesting events, but rather patterns of cause–effect relationships of the type that lend themselves to analysis and hypothesis in the forms with which we are familiar in the physical sciences. One hypothesis is that information transfer under conditions of sensory shielding is mediated by extremely low-frequency (ELF) electromagnetic waves, a proposal that does not seem to be ruled out by any obvious physical or biological facts. Further, the development of information theory makes it possible to characterize and quantify the performance of a communications channel regardless of the underlying mechanism.

For the past three years, we have had a program in the Electronics and Bioengineering Laboratory of the Stanford Research Institute (SRI) to investigate those facets of human perception that appear to fall outside the range of well-understood perceptual/processing capabilities. Of particular interest is a human information-accessing capability that we call "remote viewing." This phenomenon pertains to the ability of certain individuals to access and describe, by means of mental processes, information sources blocked from ordinary perception, and generally accepted as secure against such access.

In particular, the phenomenon we have investigated most extensively is the ability of a subject to view remote geographical locations up to several thousand kilometers distant from his physical location (given only a known person on whom to target).[1] We have carried out more than fifty experiments under controlled laboratory conditions with several individuals whose remote perceptual abilities have been developed sufficiently to allow them at times to describe correctly—often in great detail—geographical or technical material such as buildings, roads, laboratory apparatus, and the like.

As observed in the laboratory, the basic phenomenon appears to cover a range of subjective experiences variously referred to

Manuscript received July 25, 1975; revised November 7, 1975. The submission of this paper was encouraged after review of an advance proposal. This work was supported by the Foundation for Parasensory Investigation and the Parapsychology Foundation, New York, NY; the Institute of Noetic Sciences, Palo Alto, CA; and the National Aeronautics and Space Administration, under Contract NAS 7-100.
The authors are with the Electronics and Bioengineering Laboratory, Stanford Research Institute, Menlo Park, CA 94025.

[1] Our initial work in this area was reported in *Nature* [4], and reprinted in the *IEEE Commun. Soc. Newsletter*, vol. 13, Jan. 1975.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

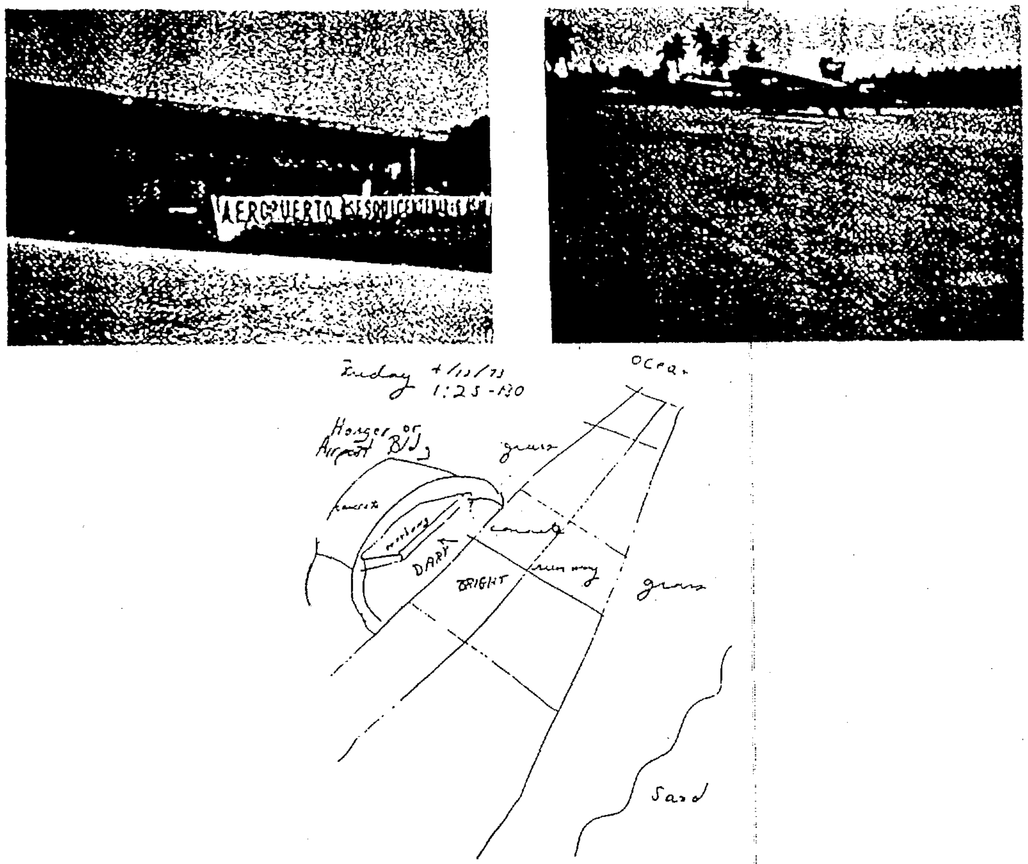


Fig. 1. Airport in San Andres, Colombia, used as remote-viewing target, along with sketch produced by subject in California.

in the literature as autoscopy (in the medical literature); exteriorization or disassociation (psychological literature); simple clairvoyance, traveling clairvoyance, or out-of-body experience (parapsychological literature); or astral projection (occult literature). We choose the term "remote viewing" as a neutral descriptive term free from prior associations and bias as to mechanisms.

The development at SRI of a successful experimental procedure to elicit this capability has evolved to the point where persons such as visiting government scientists and contract monitors, with no previous exposure to such concepts, have learned to perform well; and subjects who have trained over a one-year period have performed excellently under a variety of experimental conditions. Our accumulated data thus indicate that both specially selected and unselected persons can be assisted in developing remote perceptual abilities up to a level of useful information transfer.

In experiments of this type, we have three principal findings. First, we have established that it is possible to obtain significant amounts of accurate descriptive information about remote locations. Second, an increase in the distance from a few meters up to 4000 km separating the subject from the scene to be perceived does not in any apparent way degrade the quality or accuracy of perception. Finally, the use of Faraday cage electrical shielding does not prevent high-quality descriptions from being obtained.

To build a coherent theory for the explanation of these phenomena, it is necessary to have a clear understanding of what constitutes the phenomena. In this paper, we first briefly summarize previous efforts in this field in Section II. We then present in Sections III and IV the results of a series of more than fifty experiments with nine subjects carried out in our own laboratory, which represent a sufficiently stable data base to permit testing of various hypotheses concerning the functioning of this channel. Finally, in Section V, we indicate those areas of physics and information theory that appear to be relevant to an understanding of certain aspects of the phenomena.

First, however, we present an illustrative example generated in an early pilot experiment. As will be clear from our later discussion, this is not a "best-ever" example, but rather a typical sample of the level of proficiency that can be reached and that we have come to expect in our research.

Three subjects participated in a long-distance experiment focusing on a series of targets in Costa Rica. These subjects said they had never been to Costa Rica. In this experiment, one of the experimenters (Dr. Puthoff) spent ten days traveling through Costa Rica on a combination business/pleasure trip. This information was all that was known to the subjects about the traveler's itinerary. The experiment called for Dr. Puthoff to keep a detailed record of his location and activities, including photographs of each of seven target days at 1330 PDT. A total of twelve daily descriptions were collected before the traveler's return: six responses from one subject, five from another, and one from a third.

The third subject who submitted the single response supplied a drawing for a day in the middle of the series. (The subject's response, together with the photographs taken at the site, are shown in Fig. 1). Although Costa Rica is a mountainous country, the subject unexpectedly perceived the traveler at a beach and ocean setting. With some misgiving, he described an airport on a sandy beach and an airstrip with the ocean at the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

end (correct). An airport building also was drawn, and shown to have a large rectangular overhang (correct). The traveler had taken an unplanned one-day side trip to an offshore island and at the time of the experiment had just disembarked from a plane at a small island airport as described by the subject 4000 km away. The sole discrepancy was that the subject's drawing showed a Quonset-hut type of building in place of the rectangular structure.

The above description was chosen as an example to illustrate a major point observed a number of times throughout the program to be described. Contrary to what may be expected, a subject's description does not necessarily portray what may reasonably be expected to be correct (an educated or "safe" guess), but often runs counter even to the subject's own expectations.

We wish to stress again that a result such as the above is not unusual. The remaining submissions in this experiment provided further examples of excellent correspondences between target and response. (A target period of poolside relaxation was identified; a drive through a tropical forest at the base of a truncated volcano was described as a drive through a jungle below a large bare table mountain; a hotel-room target description, including such details as rug color, was correct; and so on.) So as to determine whether such matches were simply fortuitous—that is, could reasonably be expected on the basis of chance alone—Dr. Puthoff was asked after he had returned to blind match the twelve descriptions to his seven target locations. On the basis of this conservative evaluation procedure, which vastly underestimates the statistical significance of the individual descriptions, five correct matches were obtained. This number of matches is significant at $p = 0.02$ by exact binomial calculation.[2]

The observation of such unexpectedly high-quality descriptions early in our program led to a large-scale study of the phenomenon at SRI under secure double-blind conditions (i.e., target unknown to experimenters as well as subjects), with independent random target selection and blind judging. The results, presented in Sections III and IV, provide strong evidence for the robustness of this phenomenon whereby a human perceptual modality of extreme sensitivity can detect complex remote stimuli.

II. Background

Although we are approaching the study of these phenomena as physicists, it is not yet possible to separate ourselves entirely from the language of the nineteenth century when the laboratory study of the paranormal was begun. Consequently, we continue to use terms such as "paranormal," "telepathy," and the like. However, we intend only to indicate a process of information transfer under conditions generally accepted as secure against such transfer and with no prejudice or occult assumptions as to the mechanisms involved. As in any other scientific pursuit, the purpose is to collect the observables that result from experiments and to try to determine the functional relationships between these observables and the laws of physics as they are currently understood.

[2] The probability of a correct daily match by chance for any given transcript is $p = \frac{1}{7}$. Therefore, the probability of at least five correct matches by chance out of twelve tries can be calculated from

$$p = \sum_{t=5}^{12} \frac{12!}{t!\,(12-t)!} \left(\frac{1}{7}\right)^t \left(\frac{6}{7}\right)^{(12-t)} = 0.02.$$

Organized research into so-called psychic functioning began roughly in the time of J. J. Thomson, Sir Oliver Lodge, and Sir William Crookes, all of whom took part in the founding of the Society for Psychical Research (SPR) in 1882 in England. Crookes, for example, carried out his principal investigations with D. D. Home, a Scotsman who grew up in America and returned to England in 1855 [3]. According to the notebooks and published reports of Crookes, Home had demonstrated the ability to cause objects to move without touching them. We should note in passing that, Home, unlike most subjects, worked only in the light and spoke out in the strongest possible terms against the darkened seance rooms popular at the time [5].

Sir William Crookes was a pioneer in the study of electrical discharge in gases and in the development of vacuum tubes, some types of which still bear his name. Although everything Crookes said about electron beams and plasmas was accepted, nothing he said about the achievements of D. D. Home ever achieved that status. Many of his colleagues, who had not observed the experiments with Home, stated publicly that they thought Crookes had been deceived, to which Crookes angrily responded:

> Will not my critics give me credit for some amount of common sense? Do they not imagine that the obvious precautions, which occur to them as soon as they sit down to pick holes in my experiments, have occurred to me also in the course of my prolonged and patient investigation? The answer to this, as to all other objections is, prove it to be an error, by showing where the error lies, or if a trick, by showing how the trick is performed. Try the experiment fully and fairly. If then fraud be found, expose it; if it be a truth, proclaim it. This is the only scientific procedure, and it is that I propose steadily to pursue [3].

In the United States, scientific interest in the paranormal was centered in the universities. In 1912, John Coover [6] was established in the endowed Chair of Psychical Research at Stanford University. In the 1920's, Harvard University set up research programs with George Estabrooks and L. T. Troland [7], [8]. It was in this framework that, in 1930, William McDougall invited Dr. J. B. Rhine and Dr. Louisa Rhine to join the Psychology Department at Duke University [9]. For more than 30 years, significant work was carried out at Rhine's Duke University Laboratory. To examine the existence of paranormal perception, he used the now-famous ESP cards containing a boldly printed picture of a star, cross, square, circle, or wavy lines. Subjects were asked to name the order of these cards in a freshly shuffled deck of twenty-five such cards. To test for telepathy, an experimenter would look at the cards one at a time, and a subject suitably separated from the sender would attempt to determine which card was being viewed.

Dr. J. B. Rhine together with Dr. J. G. Pratt carried out thousands of experiments of this type under widely varying conditions [10]. The statistical results from these experiments indicated that some individuals did indeed possess a paranormal perceptual ability in that it was possible to obtain an arbitrarily high degree of improbability by continued testing of a gifted subject.

The work of Rhine has been challenged on many grounds, however, including accusations of improper handling of statistics, error, and fraud. With regard to the statistics, the general consensus of statisticians today is that if fault is to be found in Rhine's work, it would have to be on other than statistical grounds [11]. With regard to the accusations of fraud, the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

most celebrated case of criticism of Rhine's work, that of G. R. Price [12], ended 17 years after it began when the accusation of fraud was retracted by its author in an article entitled "Apology to Rhine and Soal," published in the same journal in which it was first put forward [13]. It should also be noted that parapsychological researchers themselves recently exposed fraud in their own laboratory when they encountered it [14].

At the end of the 1940's, Prof. S. G. Soal, an English mathematician working with the SPR, had carried out hundreds of card guessing experiments involving tens of thousands of calls [15]. Many of these experiments were carried out over extended distances. One of the most notable experiments was conducted with Mrs. Gloria Stewart between London and Antwerp. This experiment gave results whose probability of occurring by chance were less than 10^{-8}. With the publication of *Modern Experiments in Telepathy* by Soal and Bateman (both of whom were statisticians), it appeared that card guessing experiments produced significant results, on the average.[3]

The most severe criticism of all this work, a criticism difficult to defend against in principle, is that leveled by the well-known British parapsychological critic C. E. M. Hansel [17], who began his examination of the ESP hypothesis with the stated assumption, "In view of the *a priori* arguments against it *we know in advance* that telepathy, etc., cannot occur." Therefore, based on the "*a priori* unlikelihood" of ESP, Hansel's examination of the literature centered primarily on the possibility of fraud by subjects or investigators. He reviewed in depth four experiments which he regarded as providing the best evidence of ESP: the Pearce-Pratt distance series [18]; the Pratt-Woodruff [19] series, both conducted at Duke; and Soal's work with Mrs. Stewart and Basil Shackleton [15], as well as a more recent series by Soal and Bowden [20]. Hansel showed, in each case, how fraud *could* have been committed (by the experimenters in the Pratt-Woodruff and Soal-Bateman series, or by the subjects in the Pearce-Pratt and Soal-Bowden experiments). He gave no direct evidence that fraud *was* committed in these experiments, but said, "If the result could have arisen through a trick, the experiment must be considered unsatisfactory proof of ESP, *whether or not it is finally decided that such a trick was in fact used*" [17, p. 18]. As discussed by Honorton in a review of the field [21], Hansel's conclusion after 241 pages of careful scrutiny therefore was that these experiments were not "fraud-proof" and therefore in principle could not serve as conclusive proof of ESP.

Even among the supporters of ESP research and its results, there remained the consistent problem that many successful subjects eventually lost their ability and their scores gradually drifted toward chance results. This decline effect in no way erased their previous astronomical success; but it was a disappointment since if paranormal perception is a natural ability, one would like to see subjects improving with practice rather than getting worse.

One of the first successful attempts to overcome the decline effect was in Czechoslovakia in the work of Dr. Milan Ryzl, a chemist with the Institute of Biology of the Czechoslovakian Academy of Science and also an amateur hypnotist [22]. Through the use of hypnosis, together with feedback and reinforcement, he developed several outstanding subjects, one of whom, Pavel Stepanek, has worked with experimenters around the world for more than 10 years.

Ryzl's pioneering work came as an answer to the questions raised by the 1956 CIBA Foundation conference on extrasensory perception. The CIBA Chemical Company has annual meetings on topics of biological and chemical interest, and that same year they assembled several prominent parapsychologists to have a state-of-the-art conference on ESP [23]. The conference concluded that little progress would be made in parapsychology research until a repeatable experiment could be found; namely, an experiment that different experimenters could repeat at will and that would reliably yield a statistically significant result.

Ryzl had by 1962 accomplished that goal. His primary contribution was a decision to interact with the subject as a person, to try to build up his confidence and ability. His protocol depended on "working with" rather than "running" his subjects. Ryzl's star subject, Pavel Stepanek, has produced highly significant results with many contemporary researchers [24]-[29]. In these experiments, he was able to tell with 60-percent reliability whether a hidden card was green side or white side up, yielding statistics of a million to one with only a thousand trials.

As significant as such results are statistically, the information channel is imperfect, containing noise along with the signal. When considering how best to use such a channel, one is led to the communication theory concept of the introduction of redundancy as a means of coding a message to combat the effects of a noisy channel [30]. A prototype experiment by Ryzl using such techniques has proved to be successful. Ryzl had an assistant select randomly five groups of three digits each. These 15 digits were then encoded into binary form and translated into a sequence of green and white cards in sealed envelopes. By means of repeated calling and an elaborate majority vote protocol, Ryzl was able after 19 350 calls by Stepanek (averaging 9 s per call) to correctly identify all 15 numbers, a result significant at $p = 10^{-15}$ The hit rate for individual calls was 61.9 percent, 11 978 hits, and 7372 misses [31].

Note Added in Proof: It has been brought to our attention that a similar procedure was recently used to transmit without error the word "peace" in International Morse Code (J. C. Carpenter, "Toward the effective utilization of enhanced weak-signal ESP effects," presented at the Annual Meeting of the American Association for the Advancement of Science, New York, NY, Jan. 27, 1975).

The characteristics of such a channel can be specified in accordance with the precepts of communication theory. The bit rate associated with the information channel is calculated from [30]

$$R = H(x) - H_y(x) \tag{1}$$

where $H(x)$ is the uncertainty of the source message containing symbols with *a priori* probability p_i:

$$H(x) = -\sum_{i=1}^{2} p_i \log_2 p_i \tag{2}$$

and $H_y(x)$ is the conditional entropy based on the *a posteriori* probabilities that a received signal was actually transmitted:

$$H_y(x) = -\sum_{i,j=1}^{2} p(i,j) \log_2 p_i(j). \tag{3}$$

[3]Recently, some of the early Soal experiments have been criticized [16]. However, his long-distance experiments cited here were judged in a double-blind fashion of the type that escaped the criticism of the early experiments.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

For Stepanek's run, with $p_i = \frac{1}{2}$, $p_j(j) = 0.619$, and an average time of 9 s per choice, we have a source uncertainty $H(x) = 1$ bit and a calculated bit rate

$$R \approx 0.041 \text{ bit/symbol}$$

or

$$R/T \approx 0.0046 \text{ bit/s}.$$

(Since the 15-digit number (49.8 bits) actually was transmitted at the rate of 2.9×10^{-4} bit/s, an increase in bit rate by a factor of about 20 could be expected on the basis of a coding scheme more optimum than that used in the experiments. See, for example, Appendix A.)

Dr. Charles Tart at the University of California has written extensively on the so-called decline effect. He considers that having subjects attempt to guess cards, or perform any other repetitious task for which they receive no feedback, follows the classical technique for deconditioning any response. He thus considers card guessing "a technique for extinguishing psychic functioning in the laboratory" [32].

Tart's injunctions of the mid-sixties were being heeded at Maimonides Hospital, Brooklyn, NY, by a team of researchers that included Dr. Montague Ullman, who was director of research for the hospital; Dr. Stanley Krippner; and, later, Charles Honorton. These three worked together for several years on experiments on the occurrence of telepathy in dreams. In the course of a half-dozen experimental series, they found in their week-long sessions a number of subjects who had dreams that consistently were highly descriptive of pictorial material that a remote sender was looking at throughout the night. This work is described in detail in the experimenters' book *Dream Telepathy* [33]. Honorton is continuing work of this free-response type in which the subject has no preconceived idea as to what the target may be.

In his more recent work with subjects in the waking state, Honorton is providing homogeneous stimulation to the subject who is to describe color slides viewed by another person in a remote room. In this new work, the subject listens to white noise via earphones and views an homogeneous visual field imposed through the use of Ping-Pong ball halves to cover the subject's eyes in conjunction with diffuse ambient illumination. In this so-called Ganzfeld setting, subjects are again able, now in the waking state, to give correct and often highly accurate descriptions of the material being viewed by the sender [34].

In Honorton's work and elsewhere, it apparently has been the step away from the repetitive forced-choice experiment that has opened the way for a wide variety of ordinary people to demonstrate significant functioning in the laboratory, without being bored into a decline effect.

This survey would be incomplete if we did not indicate certain aspects of the current state of research in the USSR. It is clear from translated documents and other sources [35] that many laboratories in the USSR are engaged in paranormal research.

Since the 1930's, in the laboratory of L. Vasiliev (Leningrad Institute for Brain Research), there has been an interest in the use of telepathy as a method of influencing the behavior of a person at a distance. In Vasiliev's book *Experiments in Mental Suggestion*, he makes it very clear that the bulk of his laboratory's experiments were aimed at long-distance communication combined with a form of behavior modification; for example, putting people at a distance to sleep through hypnosis [36].

Similar behavior modification types of experiments have been carried out in recent times by I. M. Kogan, Chairman of the Bioinformation Section of the Moscow Board of the Popov Society. He is a Soviet engineer who, until 1969, published extensively on the theory of telepathic communication [37]-[40]. He was concerned with three principal kinds of experiments: mental suggestion without hypnosis over short distances, in which the percipient attempts to identify an object; mental awakening over short distances, in which a subject is awakened from a hypnotic sleep at the "beamed" suggestion from the hypnotist; and long-range (intercity) telepathic communication. Kogan's main interest has been to quantify the channel capacity of the paranormal channel. He finds that the bit rate decreases from 0.1 bit/s for laboratory experiments to 0.005 bit/s for his 1000-km intercity experiments.

In the USSR, serious consideration is given to the hypothesis that telepathy is mediated by extremely low-frequency (ELF) electromagnetic propagation. (The pros and cons of this hypothesis are discussed in Section V of this paper.) In general, the entire field of paranormal research in the USSR is part of a larger one concerned with the interaction between electromagnetic fields and living organisms [41], [42]. At the First International Congress on Parapsychology and Psychotronics in Prague, Czechoslovakia, in 1973, for example, Kholodov spoke at length about the susceptibility of living systems to extremely low-level ac and dc fields. He described conditioning effects on the behavior of fish resulting from the application of 10 to 100 μW of RF to their tank [43]. The USSR take these data seriously in that the Soviet safety requirements for steady-state microwave exposure set limits at 10 $\mu W/cm^2$, whereas the United States has set a steady-state limit of 10 mW/cm^2 [44]. Kholodov spoke also about the nonthermal effects of microwaves on animals' central nervous systems. His experiments were very carefully carried out and are characteristic of a new dimension in paranormal research.

The increasing importance of this area in Soviet research was indicated recently when the Soviet Psychological Association issued an unprecedented position paper calling on the Soviet Academy of Sciences to step up efforts in this area [45]. They recommended that the newly formed Psychological Institute within the Soviet Academy of Sciences and the Psychological Institute of the Academy of Pedagogical Sciences review the area and consider the creation of a new laboratory within one of the institutes to study persons with unusual abilities. They also recommended a comprehensive evaluation of experiments and theory by the Academy of Sciences' Institute of Biophysics and Institute for the Problems of Information Transmission.

The Soviet research, along with other behavioristically oriented work, suggests that in addition to obtaining overt responses such as verbalizations or key presses from a subject, it should be possible to obtain objective evidence of information transfer by direct measurement of physiological parameters of a subject. Kamiya, Lindsley, Pribram, Silverman, Walter, and others brought together to discuss physiological methods to detect ESP functioning, have suggested that a whole range of electroencephalogram (EEG) responses such as evoked potentials (EP's), spontaneous EEG, and the contingent negative variation (CNV) might be sensitive indicators of the detection of remote stimuli not mediated by usual sensory processes [46].

Early experimentation of this type was carried out by Douglas Dean at the Newark College of Engineering. In his

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

search for physiological correlates of information transfer, he used the plethysmograph to measure changes in the blood volume in a finger, a sensitive indicator of autonomic nervous system functioning [47]. A plethysmographic measurement was made on the finger of a subject during telepathy experiments. A sender looked at randomly selected target cards consisting of names known to the subject, together with names unknown to him (selected at random from a telephone book). The names of the known people were contributed by the subject and were to be of emotional significance to him. Dean found significant changes in the chart recording of finger blood volume when the remote sender was looking at those names known to the subject as compared with those names randomly chosen.

Three other experiments using the physiological approach have now been published. The first work by Tart [48], a later work by Lloyd [49], and most recently the work by the authors [4] all follow a similar procedure. Basically, a subject is closeted in an electrically shielded room while his EEG is recorded. Meanwhile, in another laboratory, a second person is stimulated from time to time, and the time of that stimulus is marked on the magnetic-tape recording of the subject's EEG. The subject does not know when the remote stimulus periods are as compared with the nonstimulus periods.

With regard to choice of stimulus for our own experimentation, we noted that in previous work others had attempted, without success, to detect evoked potential changes in a subject's EEG in response to a single stroboscopic flash stimulus observed by another subject [50]. In a discussion of that experiment, Kamiya suggested that because of the unknown temporal characteristics of the information channel, it might be more appropriate to use repetitive bursts of light to increase the probability of detecting information transfer [51]. Therefore, in our study we chose to use a stroboscopic flash train of 10-s duration as the remote stimulus.

In the design of the study, we assumed that the application of the remote stimulus would result in responses similar to those obtained under conditions of direct stimulation. For example, when an individual is stimulated with a low-frequency (< 30 Hz) flashing light, the EEG typically shows a decrease in the amplitude of the resting rhythm and a driving of the brain waves at the frequency of the flashes [52]. We hypothesized that if we stimulated one subject in this manner (a putative sender), the EEG of another subject in a remote room with no flash present (a receiver) might show changes in alpha (9–11 Hz) activity and possibly an EEG driving similar to that of the sender, or other coupling to the sender's EEG [53]. The receiver was seated in a visually opaque, acoustically and electrically shielded, double-walled steel room about 7 m from the sender. The details of the experiment, consisting of seven runs of thirty-six 10-s trials each (twelve periods each for 0-Hz, 6-Hz, and 16-Hz stimuli, randomly intermixed), are presented in [4]. This experiment proved to be successful. The receiver's alpha activity (9–11 Hz) showed a significant reduction in average power (−24 percent, $p < 0.04$) and peak power (−28 percent, $p < 0.03$) during 16-Hz flash stimuli as compared with periods of no-flash stimulus. [A similar response was observed for 6-Hz stimuli (−12 percent in average power, −21 percent in peak power), but the latter result did not reach statistical significance.] Fig. 2 shows an overlay of three averaged EEG spectra from one of the subject's 36 trial runs, displaying differences in alpha activity during the three stimulus conditions. Extensive control procedures were undertaken to determine if these

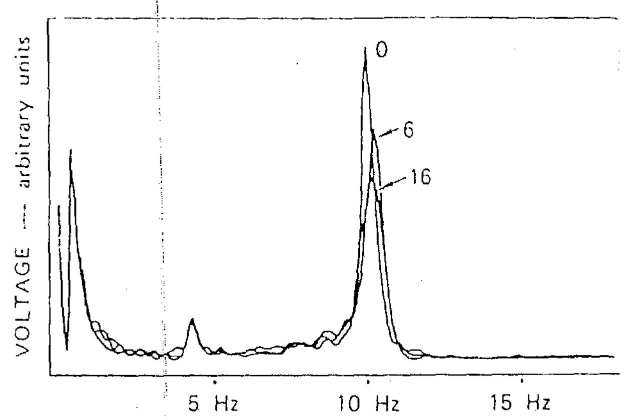


Fig. 2. Occipital EEG frequency spectra, 0–20 Hz, of one subject (H.H.) acting as receiver showing amplitude changes in the 9–11-Hz band as a function of strobe frequency. Three cases: 0-, 6-, and 16-Hz flashes (twelve trial averages).

results were produced by system artifacts, electromagnetic pickup (EMI), or subtle cueing; the results were negative [4].

As part of the experimental protocol, the subject was asked to indicate a conscious assessment for each trial (via telegraph key) as to the nature of the stimulus; analysis showed these guesses to be at chance. Thus arousal as evidenced by significant alpha blocking occurred only at the noncognitive level of physiological response. Hence the experiment provided direct physiological (EEG) evidence of perception of remote stimuli even in the absence of overt cognitive response.

Whereas in our experiments we used a remote light flash as a stimulus, Tart [48] in his work used an electrical shock to himself as sender, and Lloyd [49] simply told the sender to think of a red triangle each time a red warning light was illuminated within his view. Lloyd observed a consistent evoked potential in his subjects; whereas in our experiments and in Tart's, a reduction in amplitude and a desynchronization of alpha was observed—an arousal response. (If a subject is resting in an alpha-dominant condition and he is then stimulated, for example in any direct manner, one will observe a desynchronization and decrease in alpha power.) We consider that these combined results are evidence for the existence of noncognitive awareness of remote happenings and that they have a profound implication for paranormal research.

III. SRI Investigations of Remote Viewing

Experimentation in remote viewing began during studies carried out to investigate the abilities of a New York artist, Ingo Swann, when he expressed the opinion that the insights gained during experiments at SRI had strengthened his ability (verified in other research before he joined the SRI program) to view remote locations [54]. To test Mr. Swann's assertion, a pilot study was set up in which a series of targets from around the globe were supplied by SRI personnel to the experimenters on a double-blind basis. Mr. Swann's apparent ability to describe correctly details of buildings, roads, bridges, and the like indicated that it may be possible for a subject by means of mental imagery to access and describe randomly chosen geographical sites located several miles from the subject's position and demarcated by some appropriate means. Therefore, we set up a research program to test the remote-viewing hypothesis under rigidly controlled scientific conditions.

In carrying out this program, we concentrated on what we considered to be our principal responsibility—to resolve under unambiguous conditions the basic issue of whether or not this

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

class of paranormal perception phenomenon exists. At all times, we and others responsible for the overall program took measures to prevent sensory leakage and subliminal cueing and to prevent deception, whether intentional or unintentional. To ensure evaluations independent of belief structures of both experimenters and judges, all experiments were carried out under a protocol, described below, in which target selection at the beginning of experiments and blind judging of results at the end of experiments were handled independently of the researchers engaged in carrying out the experiments.

Six subjects, designated S1 through S6, were chosen for the study. Three were considered as gifted or experienced subjects (S1 through S3), and three were considered as learners (S4 through S6). The *a priori* dichotomy between gifted and learners was based on the experienced group having been successful in other studies conducted before this program and the learners group being inexperienced with regard to paranormal experimentation.

The study consisted of a series of double-blind tests with local targets in the San Francisco Bay Area so that several independent judges could visit the sites to establish documentation. The protocol was to closet the subject with an experimenter at SRI and at an agreed-on time to obtain from the subject a description of an undisclosed remote site being visited by a target team. In each of the experiments, one of the six program subjects served as remote-viewing subject, and SRI experimenters served as a target demarcation team at the remote location chosen in a double-blind protocol as follows.

In each experiment, SRI management randomly chose a target location from a list of targets within a 30-min driving time from SRI; the target location selected was kept blind to subject and experimenters. The target pool consisted of more than 100 target locations chosen from a target-rich environment. (Before the experimental series began, the Director of the Information Science and Engineering Division, not otherwise associated with the experiment, established the set of locations as the target pool which remained known only to him. The target locations were printed on cards sealed in envelopes and kept in the SRI Division office safe. They were available only with the personal assistance of the Division Director who issued a single random-number selected target card that constituted the traveling orders for that experiment.)

In detail: To begin the experiment, the subject was closeted with an experimenter at SRI to wait 30 min before beginning a narrative description of the remote location. A second experimenter then obtained from the Division Director a target location from a set of traveling orders previously prepared and randomized by the Director and kept under his control. The target demarcation team, consisting of two to four SRI experimenters, then proceeded by automobile directly to the target without any communication with the subject or experimenter remaining behind. The experimenter remaining with the subject at SRI was kept ignorant of both the particular target and the target pool so as to eliminate the possibility of cueing (overt or subliminal) and to allow him freedom in questioning the subject to clarify his descriptions. The demarcation team remained at the target site for an agreed-on 15-min period following the 30 min allotted for travel.[4] During the observation period, the remote-viewing subject was asked to describe his impressions of the target site into a tape recorder and to make any drawings he thought appropriate. An informal comparison was then made when the demarcation team returned, and the subject was taken to the site to provide feedback.

A. Subject S1: Experienced

To begin the series, Pat Price, a former California police commissioner and city councilman, participated as a subject in nine experiments. In general, Price's ability to describe correctly buildings, docks, roads, gardens, and the like, including structural materials, color, ambience, and activity—often in great detail—indicated the functioning of a remote perceptual ability. A Hoover Tower target, for example, was recognized and named by name. Nonetheless, in general, the descriptions contained inaccuracies as well as correct statements. A typical example is indicated by the subject's drawing shown in Fig. 3 in which he correctly described a park-like area containing two pools of water: one rectangular, 60 by 89 ft (actual dimensions 75 by 100 ft); the other circular, diameter 120 ft (actual diameter 110 ft). He incorrectly indicated the function, however, as water filtration rather than recreational swimming. (We often observe essentially correct descriptions of basic elements and patterns coupled with incomplete or erroneous analysis of function.) As can be seen from his drawing, he also included some elements, such as the tanks shown in the upper right, that are not present at the target site. We also note an apparent left–right reversal, often observed in paranormal perception experiments.

To obtain a numerical evaluation of the accuracy of the remote-viewing experiment, the experimental results were subjected to independent judging on a blind basis by an SRI research analyst not otherwise associated with the research. The subject's response packets, which contained the nine typed unedited transcripts of the tape-recorded narratives along with any associated drawings, were unlabeled and presented in random order. While standing at each target location, visited in turn, the judge was required to blind rank order the nine packets on a scale 1 to 9 (best to worst match). The statistic of interest is the sum of ranks assigned to the target-associated transcripts, lower values indicating better matches. For nine targets, the sum of ranks could range from nine to eighty-one. The probability that a given sum of ranks s or less will occur by chance is given by [55]

$$\Pr(s \text{ or less}) = \frac{1}{N^n} \sum_{i=n}^{s} \sum_{l=0}^{k} (-1)^l \binom{n}{l} \binom{i - Nl - 1}{n - 1}$$

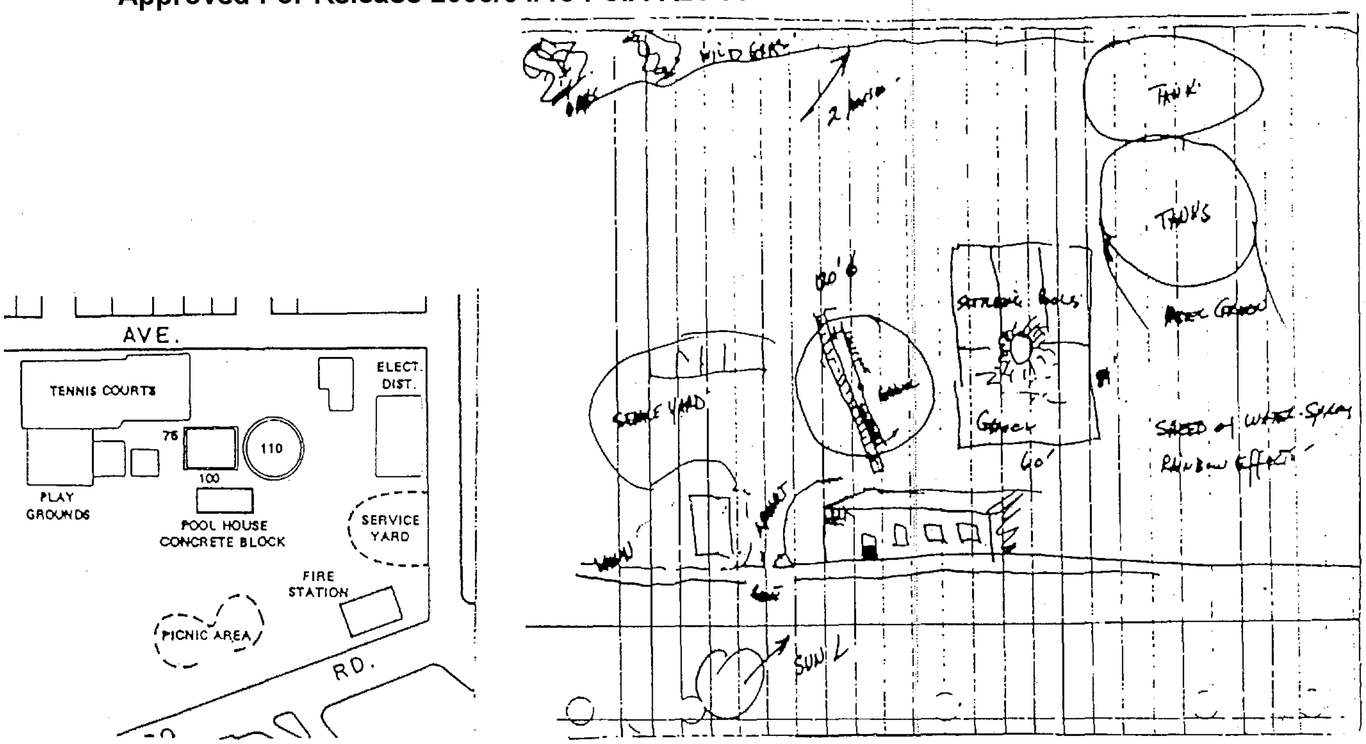
where s is obtained sum of ranks, N is number of assignable ranks, n is number of occasions on which rankings were made, and l takes on values from zero to the least positive integer k in $(i - n)/n$. (Table I is a table to enable easy application of the above formula to those cases in which $N = n$.) The sum in this case, which included seven direct hits out of the nine, was 16 (see Table II), a result significant at $p = 2.9 \times 10^{-5}$ by exact calculation.

In Experiments 3, 4, and 6 through 9, the subject was secured in a double-walled copper-screen Faraday cage. The Faraday cage provides 120-dB attenuation for plane-wave radio-frequency radiation over a range of 15 kHz to 1 GHz. For magnetic fields, the attenuation is 68 dB at 15 kHz and decreases to 3 dB at 60 Hz. The results of rank order judging (Table II) indicate that the use of Faraday cage electrical

[4]The first subject (S1) was allowed 30 min for his descriptions, but it was found that he fatigued and had little comment after the first 15 min. The viewing time was therefore reduced to 15 min for subjects S2 through S6.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4



(a) (b)

Fig. 3. Swimming pool complex as remote-viewing target. (a) City map of target location. (b) Drawing by Price (S1).

TABLE I
CRITICAL VALUES OF SUMS OF RANKS FOR PREFERENTIAL MATCHING

Number of Assignable Ranks (N)	Probability (one-tailed) that the Indicated Sum of Ranks or Less Would Occur by Chance													
	0.20	0.10	0.05	0.04	0.025	0.01	0.005	0.002	0.001	0.0005	10^{-4}	10^{-5}	10^{-6}	10^{-7}
4	7	6	5	5	5	4	4							
5	11	10	9	8	8	7	6	6	5	5				
6	16	15	13	13	12	11	10	9	8	7	6			
7	22	20	18	18	17	15	14	12	12	11	9	8		
8	29	27	24	24	22	20	19	17	16	15	13	11	9	8
9	37	34	31	30	29	26	24	22	21	20	17	14	12	10
10	46	42	39	38	36	33	31	29	27	25	22	19	16	13
11	56	51	48	47	45	41	38	36	34	32	28	24	20	17
12	67	61	58	56	54	49	47	43	41	39	35	30	25	22

Note: This table applies only to those special cases in which the number of occasions on which objects are being ranked (n) is equal to the number of assignable ranks (N). Each entry represents the largest number that is significant at the indicated p-level. *Source:* R. L. Morris [55].

shielding does not prevent high-quality descriptions from being obtained.

As a backup judging procedure, a panel of five additional SRI scientists not otherwise associated with the research were asked simply to blind match the unedited typed transcripts (with associated drawings) generated by the remote viewer against the nine target locations which they independently visited in turn. The transcripts were unlabeled and presented in random order. A correct match consisted of a transcript of a given date being matched to the target of that date. Instead of the expected number of 1 match each per judge, the number of correct matches obtained by the five judges was 7, 6, 5, 3, and 3, respectively. Thus, rather than the expected total number of 5 correct matches from the judges, 24 such matches were obtained.

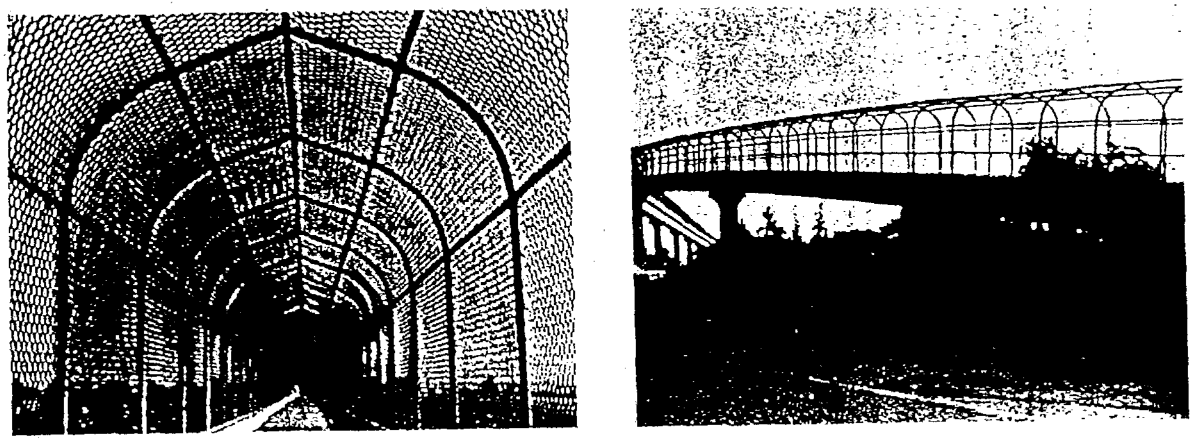
B. Subject S4: Learner

This experiment was designed to be a replication of our previous experiment with Price, the first replication attempted. The subject for this experiment was Mrs. Hella Hammid, a gifted professional photographer. She was selected for this series on the basis of her successful performance as a percipient in the EEG experiment described earlier. Outside of that interaction, she had no previous experience with apparent paranormal functioning.

At the time we began working with Mrs. Hammid, she had no strong feelings about the likelihood of her ability to succeed in this task. This was in contrast to both Ingo Swann who had come to our laboratory fresh from a lengthy and apparently successful series of experiments with Dr. Gertrude Schmeidler at City College of New York [56] and Pat Price

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4



PEDESTRIAN OVERPASS TARGET

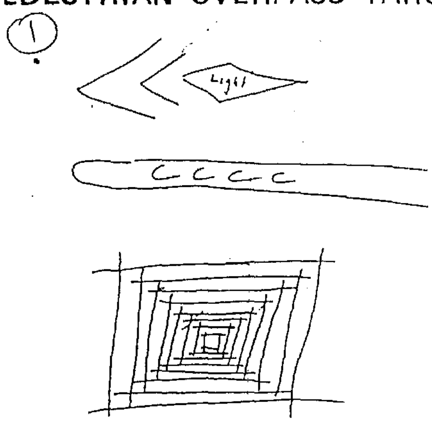


Fig. 4. Subject Hammid (S4) drawing, described as "some kind of diagonal trough up in the air."

TABLE II
DISTRIBUTION OF RANKINGS ASSIGNED TO TRANSCRIPTS ASSOCIATED WITH EACH TARGET LOCATION FOR EXPERIENCED SUBJECT PRICE (S1)

Target Location	Distance (km)	Rank of Associated Transcript
Hoover Tower, Stanford	3.4	1
Baylands Nature Preserve, Palo Alto	6.4	1
Radio telescope, Portola Valley	6.4	1
Marina, Redwood City	6.8	1
Bridge toll plaza, Fremont	14.5	6
Drive-in theater, Palo Alto	5.1	1
Arts and Crafts Plaza, Menlo Park	1.9	1
Catholic Church, Portola Valley	8.5	3
Swimming pool complex, Palo Alto	3.4	1
Total sum of ranks		16 ($p=2.9\times10^{-5}$)

who felt that he used his remote-viewing ability in his everyday life.

In comparison with the latter two, many people are more influenced by their environment and are reluctant under public scrutiny to attempt activities that are generally thought to be impossible. Society often provides inhibition and negative feedback to the individual who might otherwise have explored his own nonregular perceptual ability. We all share an historical tradition of "the stoning of prophets and the burning of witches" and, in more modern times, the hospitalization of those who claim to perceive things that the majority do not admit to seeing. Therefore, in addition to maintaining scientific rigor, one of our primary tasks as researchers is to provide an environment in which the subject feels safe to explore the possibility of paranormal perception. With a new subject, we also try to stress the nonuniqueness of the ability because from our experience paranormal functioning appears to be a latent ability that all subjects can articulate to some degree.

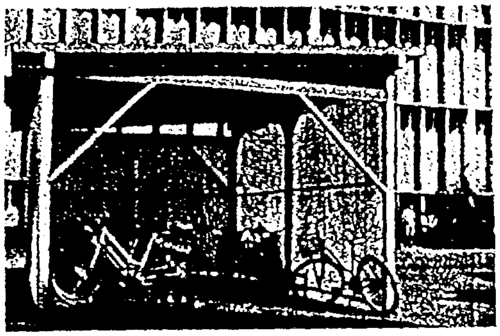
Because of Mrs. Hammid's artistic background, she was capable of drawing and describing visual images that she could not identify in any cognitive or analytic sense. When the target demarcation team went to a target location which was a pedestrian overpass, the subject said that she saw "a kind of trough up in the air," which she indicated in the upper part of her drawing in Fig. 4. She went on to explain, "If you stand where they are standing you will see something like this," indicating the nested squares at the bottom of Fig. 4. As it turned out, a judge standing where she indicated would have a view closely resembling what she had drawn, as can be seen from the accompanying photographs of the target location. It needs to be emphasized, however, that judges did not have access to our photographs of the site, used here for illustrative purposes only, but rather they proceeded to each of the target locations by list.

In another experiment, the subject described seeing "an open barnlike structure with a pitched roof." She also saw a "kind of slatted side to the structure making light and dark bars on the wall." Her drawing and a photograph of the associated bicycle shed target are shown in Fig. 5. (Subjects are encouraged to make drawings of anything they visualize and associate with the remote location because drawings they make are in general more accurate than their verbal description.)

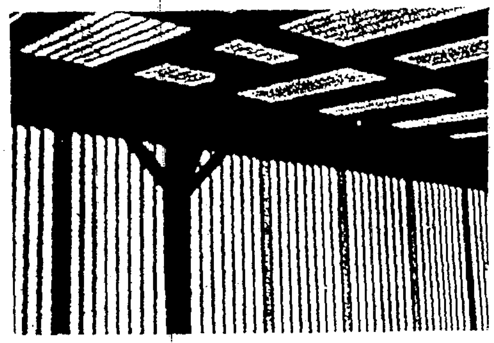
As in the original series with Price, the results of the nine-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4



BICYCLE SHED TARGET



DETAIL OF BICYCLE SHED

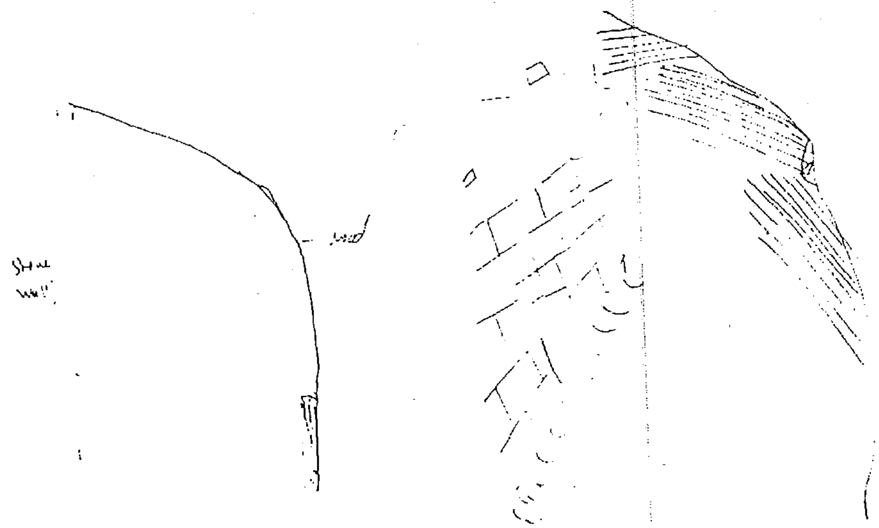


Fig. 5. Subject Hammid (S4) response to bicycle shed target described as an open "barn-like building" with "slats on the sides" and a "pitched roof."

TABLE III
DISTRIBUTION OF RANKINGS ASSIGNED TO TRANSCRIPTS ASSOCIATED WITH EACH TARGET LOCATION FOR LEARNER SUBJECT HAMMID (S4)

Target Location	Distance (km)	Rank of Associated Transcript
Methodist Church, Palo Alto	1.9	1
Ness Auditorium, Menlo Park	0.2	1
Merry-go-round, Palo Alto	3.4	1
Parking garage, Mountain View	8.1	2
SRI International Courtyard, Menlo Park	0.2	1
Bicycle shed, Menlo Park	0.1	2
Railroad trestle bridge, Palo Alto	1.3	2
Pumpkin patch, Menlo Park	1.3	1
Pedestrian overpass, Palo Alto	5.0	2
Total sum of ranks		13 ($p = 1.8 \times 10^{-6}$)

experiment series were submitted for independent judging on a blind basis by an SRI research analyst not otherwise associated with the research. While at each target location, visited in turn, the judge was required to blind rank order the nine unedited typed manuscripts of the tape-recorded narratives, along with any associated drawings generated by the remote viewer, on a scale 1 to 9 (best to worst match). The sum of ranks assigned to the target-associated transcripts in this case was 13, a result significant at $p = 1.8 \times 10^{-6}$ by exact calculation (see Table I and discussion), and included five direct hits and four second ranks (Table III).

Again, as a backup judging procedure, a panel of five additional judges not otherwise associated with the research were asked simply to blind match the unedited typed transcripts and associated drawings generated by the remote viewer, against the nine target locations which they independently visited in turn. A correct match consisted of a transcript of a given date being matched to the target of that date. Instead of the expected number of 1 match each per judge, the number of correct matches obtained by the five judges was 5, 3, 3, 2, and 2, respectively. Thus, rather than the expected total number of 5 correct matches from the judges, 15 such matches were obtained.

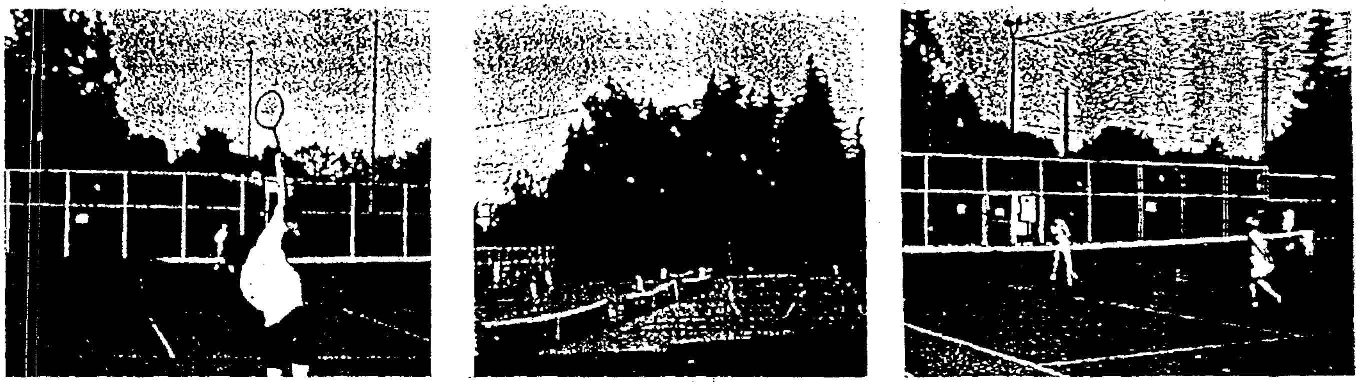
C. Subjects S2 and S3: Experienced

Having completed a series of 18 remote-viewing experiments, 9 each with experienced subject S1 (Price) and learner S4 (Hammid), additional replication experiments, four with each subject, were carried out with experienced subjects S2 (Elgin) and S3 (Swann) and learners S5 and S6. To place the judging on a basis comparable to that used with S1 and S4, the four transcripts each of experienced subjects S2 and S3 were combined into a group of eight for rank order judging to be compared with the similarly combined results of the learners S5 and S6.

The series with S2 (Elgin, an SRI research analyst) provided a further example of the dichotomy between verbal and drawing responses. (As with medical literature, case histories often are more illuminating than the summary of results.) The experiment described here was the third conducted with this

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4



TARGET—TENNIS COURTS

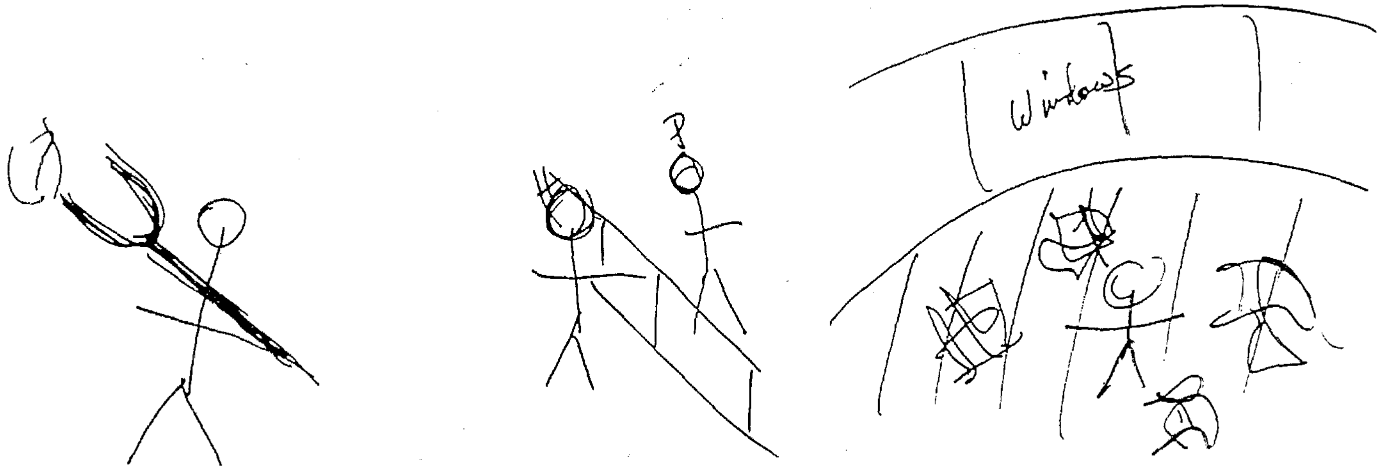


Fig. 6. Subject Elgin (S2) drawings in response to tennis court target.

subject. It was a demonstration experiment for a government visitor who had heard of our work and wanted to evaluate our experimental protocol.

In the laboratory, the subject, holding a bearing compass at arm's length, began the experiment by indicating the direction of the target demarcation team correctly to within 5°. (In all four experiments with this subject, he has always been within 10° of the correct direction in this angular assessment.) The subject then generated a 15-min tape-recorded description and the drawings shown in Fig. 6.

In discussing the drawings, Elgin indicated that he was uncertain as to the action, but had the impression that the demarcation team was located at a museum (known to him) in a particular park. In fact, the target was a tennis court located in that park about 90 m from the indicated museum. Once again, we note the characteristic (discussed earlier) of a resemblance between the target site and certain gestalt elements of the subject's response, especially in regard to the drawings, coupled with incomplete or erroneous analysis of the significances. Nonetheless, when rank ordering transcripts 1 through 8 at the site, the judge ranked this transcript as 2. This example illustrates a continuing observation that most of the correct information related to us by subjects is of a nonanalytic nature pertaining to shape, form, color, and material rather than to function or name.

A second example from this group, generated by S3 (Swann), indicates the level of proficiency that can be attained with practice. In the two years since we first started working with Swann, he has been studying the problem of separating the external signal from the internal noise. In our most recent experiments, he dictates two lists for us to record. One list contains objects that he "sees," but does not think are located at the remote scene. A second list contains objects that he thinks are at the scene. In our evaluation, he has made much progress in this most essential ability to separate memory and imagination from paranormal inputs. This is the key to bringing the remote-viewing channel to fruition with regard to its potential usefulness.

The quality of transcript that can be generated by this process is evident from the results of our most recent experiment with Swann. The target location chosen by the usual double-blind protocol was the Palo Alto City Hall. Swann described a tall building with vertical columns and "set in" windows. His sketch, together with the photograph of the site, is shown in Fig. 7. He said there was a fountain, "but I don't hear it." At the time the target team was at the City Hall during the experiment, the fountain was not running. He also made an effort to draw a replica of the designs in the pavement in front of the building, and correctly indicated the number of trees (four) in the sketch.

For the entire series of eight, four each from S2 and S3, the numerical evaluation based on blind rank ordering of transcripts at each site was significant at $p = 3.8 \times 10^{-4}$ and included three direct hits and three second ranks for the target-associated transcripts (see Table IV).

D. Subjects S5 and S6: Learners

To complete the series, four experiments each were carried out with learner subjects S5 and S6, a man and woman on the SRI professional staff. The results in this case, taken as a

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

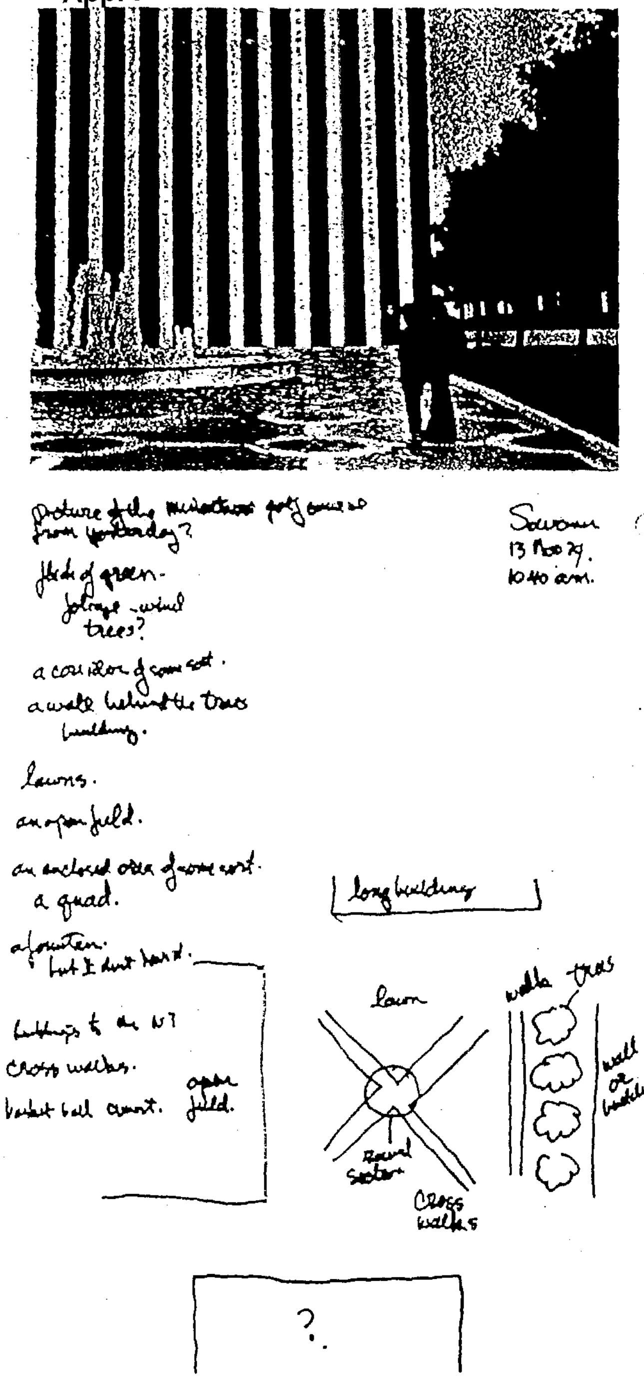


Fig. 7. Subject Swann (S3) response to City Hall target.

TABLE IV
DISTRIBUTION OF RANKINGS ASSIGNED TO TRANSCRIPTS ASSOCIATED WITH EACH TARGET LOCATION FOR EXPERIENCED SUBJECTS ELGIN (S2) AND SWANN (S3)

Subject	Target Location	Distance (km)	Rank of Associated Transcript
S2	BART Station (Transit System), Fremont	16.1	1
S2	Shielded room, SRI, Menlo Park	0.1	2
S2	Tennis court, Palo Alto	3.4	2
S2	Golf course bridge, Stanford	3.4	2
S3	City Hall, Palo Alto	2.0	1
S3	Miniature golf course, Menlo Park	3.0	1
S3	Kiosk in park, Menlo Park	0.3	3
S3	Baylands Nature Preserve, Palo Alto	6.4	3
	Total sum of ranks		15 ($p = 3.8 \times 10^{-4}$)

TABLE V
DISTRIBUTION OF RANKINGS ASSIGNED TO TRANSCRIPTS ASSOCIATED WITH EACH TARGET LOCATION FOR LEARNER SUBJECTS S5 AND S6

Subject	Target Location	Distance (km)	Rank of Associated Transcript
S5	Pedestrian overpass, Palo Alto	5.0	3
S5	Railroad trestle bridge, Palo Alto	1.3	6
S5	Windmill, Portola Valley	5.5	2
S5, S6	White Plaza, Stanford (2)	3.8	1
S6	Airport, Palo Alto	5.5	2
S6	Kiosk in Park, Menlo Park	0.3	5
S6	Boathouse, Stanford	4.0	1
	Total sum of ranks		20 ($p = 0.08$, NS)

group, did not differ significantly from chance. For the series of eight (judged as a group of seven since one target came up twice, once for each subject), the numerical evaluation based on blind rank ordering of transcripts at each site was nonsignificant at $p = 0.08$, even though there were two direct hits and two second ranks out of the seven (see Table V).

One of the direct hits, which occurred with subject S6 in her *first* experiment, provides an example of the "first-time effect" that has been rigorously explored and is well-known to experimenters in the field [57]. The outbound experimenter obtained, by random protocol from the pool, a target blind to the experimenter with the subject, as is our standard procedure, and proceeded to the location. The subject, a mathematician in the computer science laboratory who had no previous experience in remote viewing, began to describe a large square with a fountain. Four minutes into the experiment, she recognized the location and correctly identified it by name (see Fig. 8). (It should be noted that in the area from which the target locations were drawn there are other fountains as well, some of which were in the target pool.) As an example of the style of the narratives generated during remote viewing with inexperienced subjects and of the part played by the experimenter remaining with the subject in such a case, we have included the entire unedited text of this experiment as Appendix B.

E. Normal and Paranormal: Use of Unselected Subjects in Remote Viewing

After more than a year of following the experimental protocol described above and observing that even inexperienced subjects generated results better than expected, we initiated a series of experiments to explore further whether individuals other than putative "psychics" can demonstrate the remote-viewing ability. To test this idea, we have a continuing program to carry out additional experiments of the outdoor type with new subjects whom we have no *a priori* reason to believe have paranormal perceptual ability. To date we have collected data from five experiments with two individuals in this category: a man and a woman who were visiting government scientists interested in observing our experimental protocols. The motivation for these particular experiments was twofold. First, the experiments provide data that indicate the level of proficiency that can be expected from unselected volunteers.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

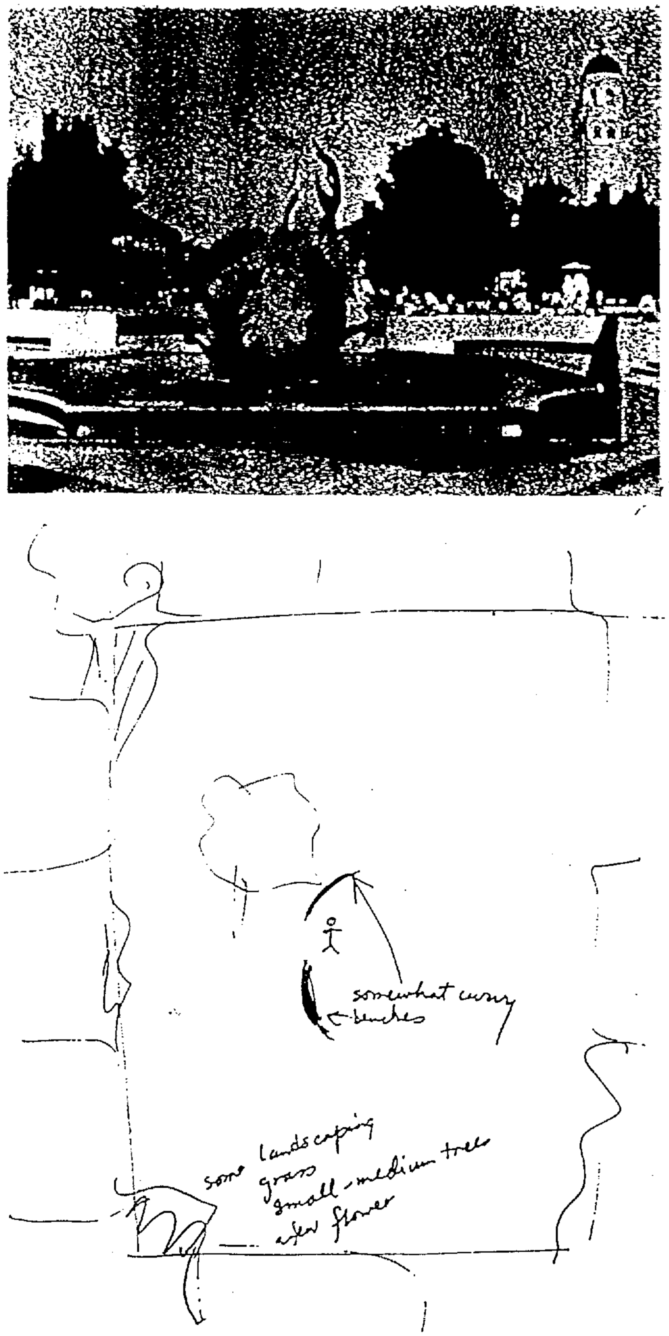


Fig. 8. Subject (S6) drawing of White Plaza, Stanford University. Subject drew what she called "curvy benches" and then announced correctly that the place was "White Plaza at Stanford."

Second, when an individual observes a successful demonstration experiment involving another person as subject, it inevitably occurs to him that perhaps chicanery is involved. We have found the most effective way to settle this issue for the observer is to have the individual himself act as a subject so as to obtain personal experience against which our reported results can be evaluated.

The first visitor (V1) was invited to participate as a subject in a three-experiment series. All three experiments contained elements descriptive of the associated target locations; the quality of response increased with practice. The third response is shown in Fig. 9, where again the pattern elements in the drawing appeared to be a closer match than the subject's analytic interpretation of the target object as a cupola.

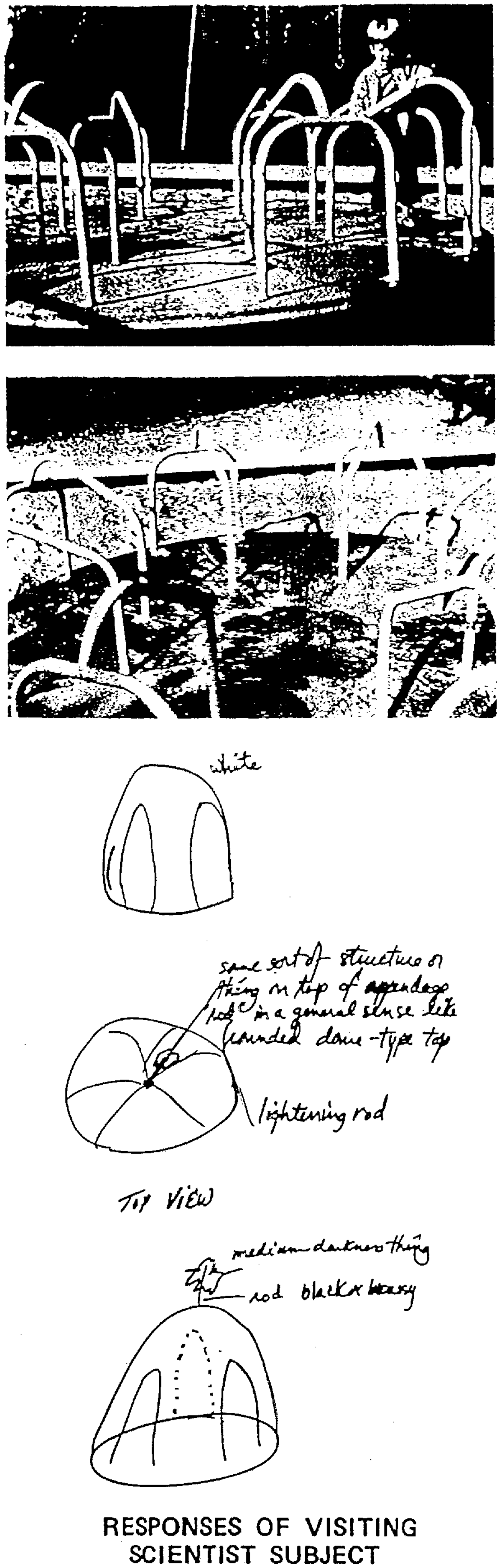


Fig. 9. Subject (V1) drawing of merry-go-round target.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

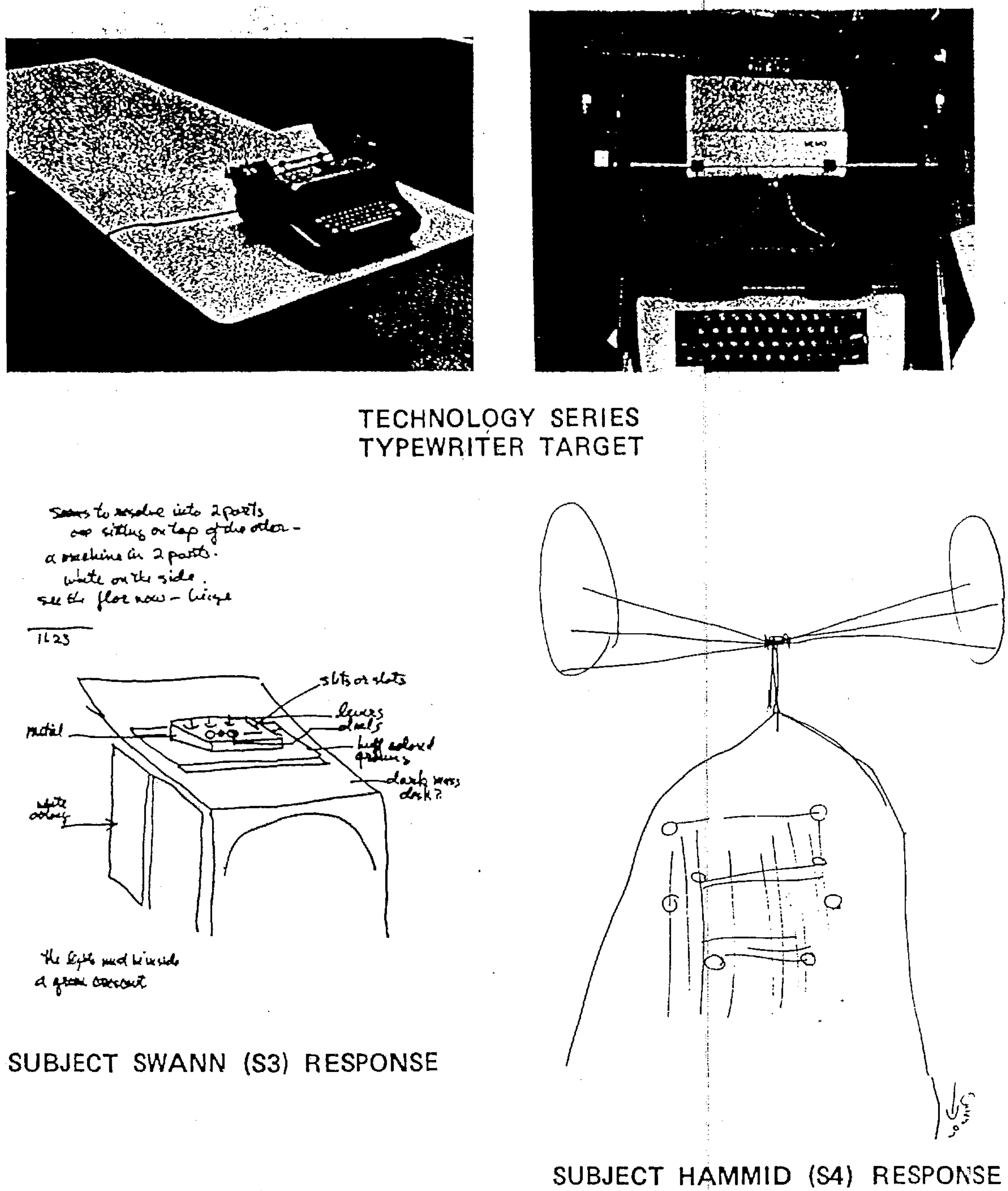


Fig. 10. Drawings of a typewriter target by two subjects.

The second visitor [V2] participated as a subject in two experiments. In his first experiment, he generated one of the higher signal-to-noise results we have observed. He began his narrative, "There is a red A-frame building and next to it is a large yellow thing [a tree—Editor]. Now further left there is another A-shape. It looks like a swing-set, but it is pushed down in a gully so I can't see the swings." [All correct.] He then went on to describe a lock on the front door that he said "looks like it's made of laminated steel, so it must be a Master lock." [Also correct.]

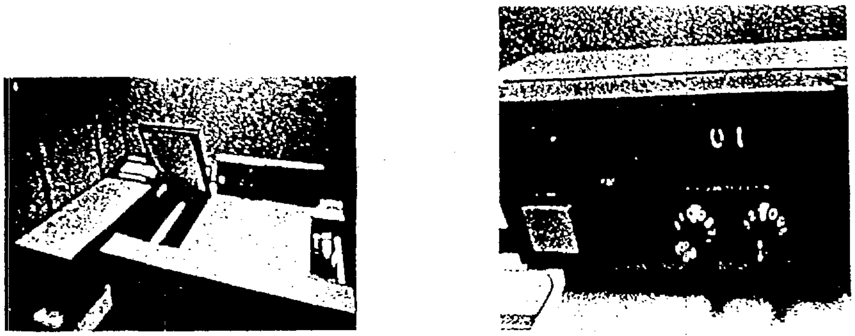
For the series of five—three from the first subject and two from the second—the numerical evaluation based on blind rank ordering of the transcripts at each site was significant at $p = 0.017$ and included three direct hits and one second-rank for the target-associated transcripts. (See Table VI.)

TABLE VI
DISTRIBUTION OF RANKINGS ASSIGNED TO TRANSCRIPTS ASSOCIATED WITH EACH TARGET LOCATION FOR VISITOR SUBJECTS V1 AND V2

Subject	Target Location	Distance (km)	Rank of Associated Transcript
V1	Bridge over stream, Menlo Park	[illegible].3	1
V1	Baylands Nature Preserve, Palo Alto	[illegible].4	2
V1	Merry-go-round, Palo Alto	[illegible].4	1
V2	Windmill, Portola Valley	[illegible].5	1
V2	Apartment swimming pool, Mountain View	[illegible].1	3
	Total sum of ranks		8 (p=0.017)

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4



TARGET LOCATION: XEROX MACHINE
(TECHNOLOGY SERIES)

TO ADD INTEREST TO TARGET
LOCATION EXPERIMENTER WITH
HIS HEAD BEING XEROXED

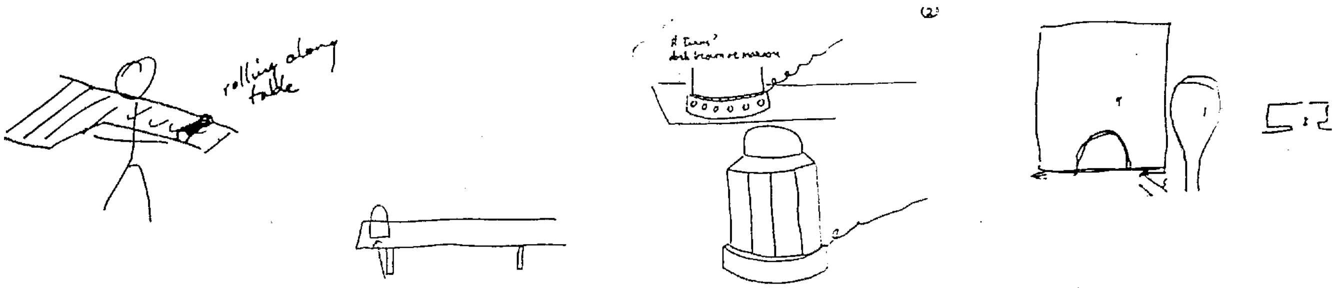


Fig. 11. Drawings by three subjects (S2, S3, and V3) for Xerox machine target. When asked to describe the square at upper left of response on the right, subject (V3) said, "There was this predominant light source which might have been a window, and a working surface which might have been the sill, or a working surface or desk." Earlier the subject had said, "I have the feeling that there is something silhouetted against the window."

Observations with unselected subjects such as those described above indicate that remote viewing may be a latent and widely distributed perceptual ability.

F. Technology Series: Short-Range Remote Viewing

Because remote viewing is a perceptual ability, we considered it important to obtain data on its resolution capabilities. To accomplish this, we turned to the use of indoor technological targets.

Twelve experiments were carried out with five different subjects, two of whom were visiting government scientists. They were told that one of the experimenters would be sent by random protocol to a laboratory within the SRI complex and that he would interact with the equipment or apparatus at that location. It was further explained that the experimenter remaining with the subject was, as usual, kept ignorant of the contents of the target pool to prevent cueing during questioning. (Unknown to subjects, targets in the pool were used with replacement; one of the goals of this particular experiment was to obtain multiple responses to a given target to investigate whether correlation of a number of subject responses would provide enhancement of the signal-to-noise ratio.) The subject was asked to describe the target both verbally (tape recorded) and by means of drawings during a time-synchronized 15-min interval in which the outbound experimenter interacted in an appropriate manner with the equipment in the target area.

In the twelve experiments, seven targets were used: a drill press, Xerox machine, video terminal, chart recorder, four-state random number generator, machine shop, and typewriter. Three of these were used twice (drill press, video terminal, and typewriter) and one (Xerox machine) came up three times in our random selection procedure.

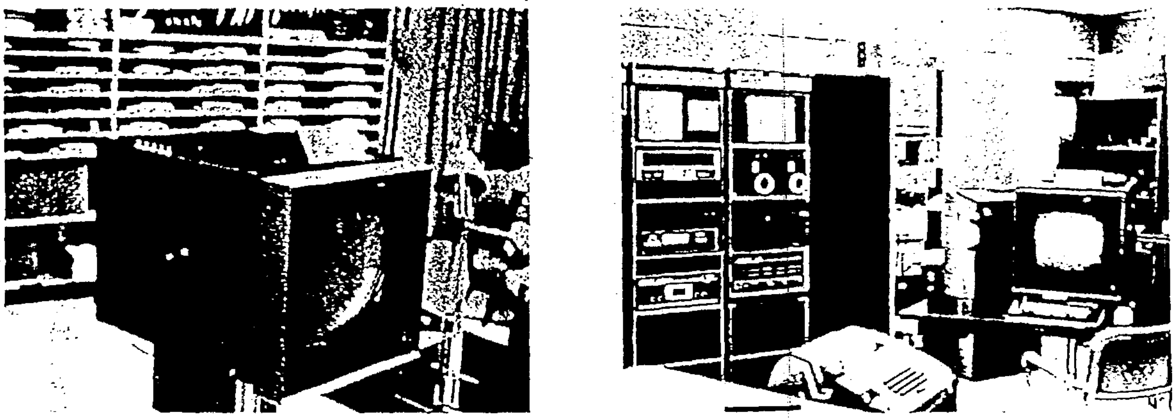
Comparisons of the targets and subject drawings for three of the multiple-response cases (the typewriter, Xerox machine, and video terminal) are shown in Figs. 10, 11, and 12. As is apparent from these illustrations alone, the experiments provide circumstantial evidence for an information channel of useful bit rate. This includes experiments in which visiting government scientists participated as subjects (Xerox machine and video terminal) to observe the protocol. In general, it appears that use of multiple-subject responses to a single target provides better signal-to-noise ratio than target identification by a single individual. This conclusion is borne out by the judging described below.

Given that in general the drawings constitute the most accurate portion of a subject's description, in the first judging procedure a judge was asked simply to blind match only the drawings (i.e., without tape transcripts) to the targets. Multiple-subject responses to a given target were stapled together, and thus seven subject-drawing response packets were to be matched to the seven different targets for which drawings were made. The judge did *not* have access to our photographs of the target locations, used for illustration purposes only, but rather proceeded to each of the target locations by list. While standing at each target location, the judge was required to rank order the seven subject-drawing response packets (presented in random order) on a scale 1 to 7 (best to worst match). For seven targets, the sum of ranks could range from 7 to 49. The sum in this case, which included 1 direct hit and 4 second ranks out of the 7 (see Table VII) was 18, a result significant at $p = 0.036$.

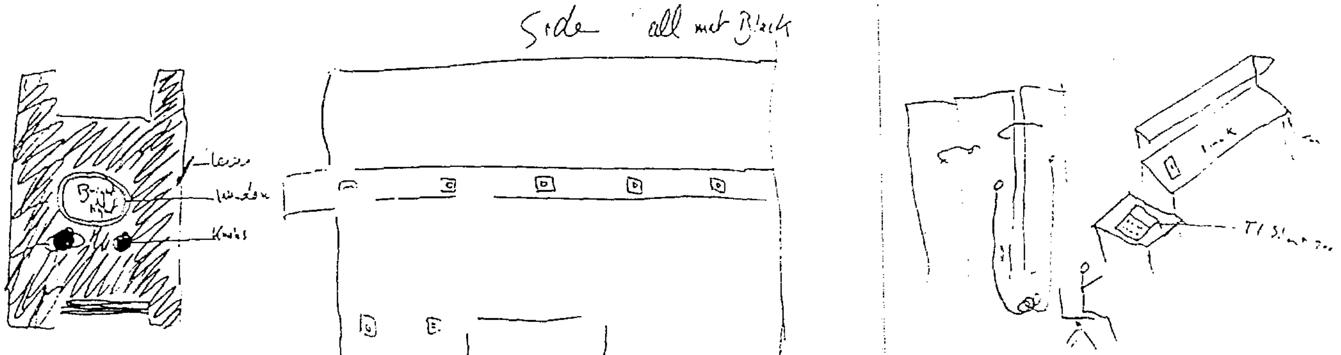
In the second more detailed effort at evaluation, a visiting scientist selected at random one of the 12 data packages (a drill press experiment), sight unseen and submitted it for independent analysis to an engineer with a request for an esti-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4



TARGET: VIDEO MONITOR FOR TEXT EDITING (TECHNOLOGY SERIES)



(a) (b)

Fig. 12. Drawing by two subjects of a video monitor target. (a) Subject (S4) drawing of "box with light coming out of it . . . painted flat black and in the middle of the room." (b) Second subject (V2) saw a computer terminal with relay racks in the background.

TABLE VII
DISTRIBUTION OF RANKINGS ASSIGNED TO SUBJECT DRAWINGS ASSOCIATED WITH EACH TARGET LOCATION

Subject	Target	Rank of Associated Drawings
S3, S4	Drill press	2
S2, S3, V3	Xerox machine	2
S4, V2	Video terminal	1
S3	Chart recorder	2
S4	Random number generator	6
S4	Machine shop	3
S3, S4	Typewriter	2
	Total sum of ranks	18 (p=0.036)

TABLE VIII
SUMMARY: REMOTE VIEWING

Subject	Number of Experiments	p-Value, Rank Order Judging
With natural targets		
S1 (experienced)	9	2.9×10^{-5}
S2 and S3 (experienced)	8	3.8×10^{-4}
S4 (learner)	9	1.8×10^{-6}
S5 and S6 (learners)	8	0.08 (NS)
V1 and V2 (learners/visitors)	5	0.017
With technology targets		
S2, S3, S4, V2, V3	12	0.036

mate as to what was being described. The analyst, blind as to the target and given only the subject's taped narrative and drawing (Fig. 13), was able, from the subject's description alone, to correctly classify the target as a "man-sized vertical boring machine."

G. Summary of Remote Viewing Results

1) Discussion: The descriptions supplied by the subjects in the experiments involving remote viewing of natural targets or laboratory apparatus, although containing inaccuracies, were sufficiently accurate to permit the judges to differentiate among various targets to the degree indicated. A summary tabulation of the statistical evaluations of these fifty-one experiments with nine subjects is presented in Table VIII. The overall result, evaluated conservatively on the basis of a judging procedure that ignores transcript quality beyond that necessary to rank order the data packets (vastly underestimating the statistical significance of individual descriptions), clearly indicates the presence of an information channel of useful bit rate. Furthermore, it appears that the principal difference between experienced subjects and inexperienced volunteers is *not* that the latter never exhibit the faculty, but rather that their results are simply less reliable, more sporadic. Nevertheless, as described earlier, individual transcripts from the inexperienced group of subjects number among some of the best obtained. Such observations indicate a hypothesis that remote viewing may be a latent and widely distributed perceptual ability.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

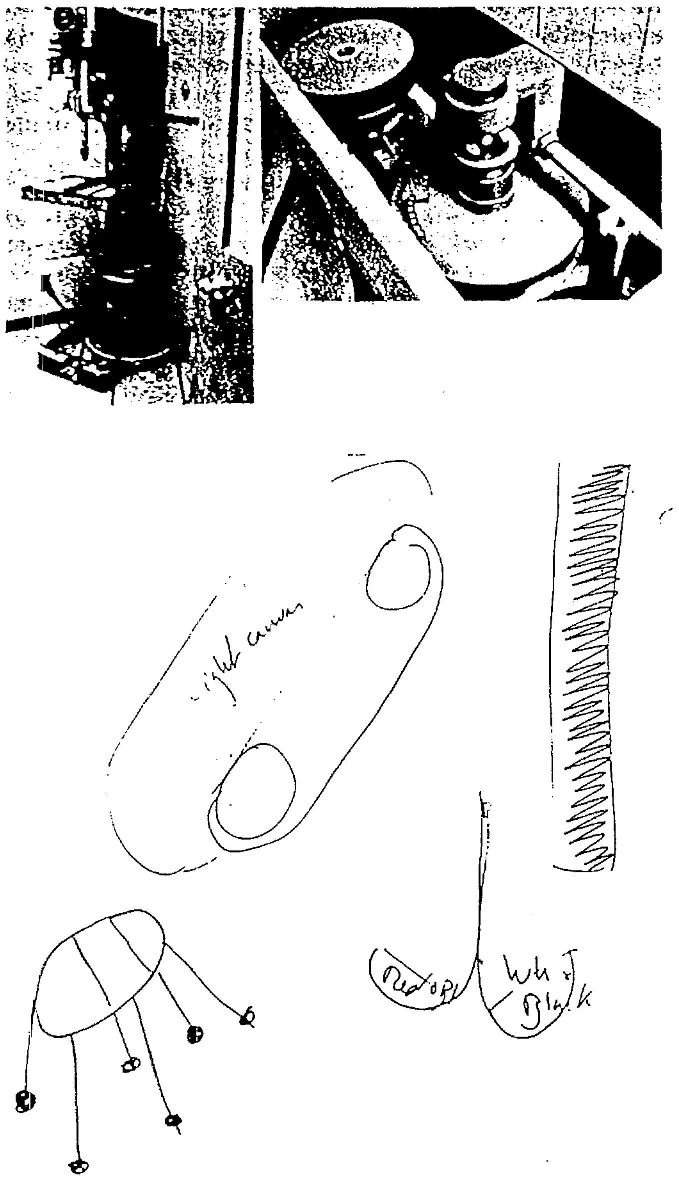


Fig. 13. Subject (S4) drawing of drill press showing belt drive, stool, and a "vertical graph that goes up and down."

Thus the primary achievement of the SRI program was the elicitation of high-quality remote viewing from individuals who agreed to act as subjects. Criticism of this claim could in principle be put forward on the basis of three potential flaws. 1) The study could involve naiveté in protocol that permits various forms of cueing, intentional or unintentional. 2) The experiments discussed could be selected out of a larger pool of experiments of which many are of poorer quality. 3) Data for the reported experiments could be edited to show only the matching elements, the nonmatching elements being discarded.

All three criticisms, however, are invalid. First, with regard to cueing, the use of double-blind protocols ensures that none of the persons in contact with the subject can be aware of the target. Second, selection of experiments for reporting did not take place; every experiment was entered as performed on a master log and is included in the statistical evaluations. Third, data associated with a given experiment remain unedited; all experiments are tape recorded and all data are included unedited in the data package to be judged and evaluated.

In the process of judging—attempting to match transcripts against targets on the basis of the information in the transcripts—some patterns and regularities in the transcript descriptions became evident, particularly regarding individual styles in remote viewing and in the perceptual form of the descriptions given by the subjects. These patterns and the judging procedure are discussed below.

a) Styles of response: The fifty-one transcripts were taken from nine different subjects. Comparing the transcripts of one subject with those of another revealed that each pattern tended to focus on certain aspects of the remote target complex and to exclude others, so that each had an individual pattern of response, like a signature.

Subject S3, for example, frequently responded with topographical descriptions, maps, and architectural features of the target locations. Subject S2 often focused on the behavior of the remote experimenter or the sequence of actions he carried out at the target. The transcripts of subject S4, more than those of other subjects, had descriptions of the feel of the location, and experiential or sensory gestalts—for example, light/dark elements in the scene and indoor/outdoor and enclosed/open distinctions. Prominent features of S1's transcripts were detailed descriptions of what the target persons were concretely experiencing, seeing, or doing—for example, standing on asphalty blacktop overlooking water; looking at a purple iris.

The range of any individual subject's responses was wide. Anyone might draw a map or describe the mood of the remote experimenter, but the consistency of each subject's overall approach suggests that just as individual descriptions of a *directly* viewed scene would differ, so these differences also occur in *remote*-viewing processes.

b) Nature of the description: The concrete descriptions that appear most commonly in transcripts are at the level of subunits of the overall scene. For example, when the target was a Xerox copy machine, the responses included (S2) a rolling object (the moving light) or dials and a cover that is lifted (S3), but the machine as a whole was not identified by name or function.

In a few transcripts, the subjects correctly identified and named the target. In the case of a computer terminal, the subject (V2) apparently perceived the terminal and the relay racks behind it. In the case of targets which were Hoover Tower and White Plaza, the subjects (S1 and S6, respectively) seemed to identify the locations through analysis of their initial images of the elements of the target.

There were also occasional incorrect identifications. Gestalts were incorrectly named; for example, swimming pools in a park were identified as water storage tanks at a water filtration plant (S1).

The most common perceptual level was thus an intermediate one—the individual elements and items that make up the target. This is suggestive of a scanning process that takes sample perceptions from within the overall environment.

When the subjects tried to make sense out of these fragmentary impressions, they often resorted to metaphors or constructed an image with a kind of perceptual inference. From a feeling of the target as an "august" and "solemn" building, a subject (S4) said it might be a library; it was a church. A pedestrian overpass above a freeway was described as a conduit (S4). A rapid transit station, elevated above the countryside, was associated with an observatory (S2). These responses seem to be the result of attempts to process partial informa-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

tion: similarly, this occurs in other parapsychological experiments. These observations are compatible with the hypotheses that information received in a putative remote-viewing mode is processed piecemeal in pattern form (consistent with a low bit rate process, but not necessarily requiring it); and the errors arise in the processes of attempted integration of the data into larger patterns directed toward verbal labeling.

When the subjects augmented the verbal transcripts with drawings or sketches, these often expressed the target elements more accurately than the verbal descriptions. Thus the drawings tended to correspond to the targets more clearly and precisely than the words of the transcript.

The descriptions given by the subjects sometimes went beyond what the remote experimenter experienced, at least consciously. For example, one subject (S4) described and drew a belt drive at the top of a drill press that was invisible even to the remote experimenter who was operating the machine; another subject (S1) described a number of items behind shrubbery and thus not visible to members of the demarcation team at the site.

Curiously, objects in motion at the remote site were rarely mentioned in the transcript. For example, trains crossing the railroad trestle target were not described, though the remote experimenter stood very close to them.

Also in a few cases, the subject descriptions were inaccurate regarding size of structures. A 20-ft courtyard separating two buildings was described as 200 ft wide, and a small shed was expanded to a barn-like structure.

c) Blind judging of transcripts: The judging procedure entailed examining the transcripts for a given experimental series and attempting to match the transcripts with the correct targets on the basis of their correspondences. The transcripts varied from coherent and accurate descriptions to mixtures of correspondences and noncorrespondences. Since the judge did not know *a priori* which elements of the descriptions were correct or incorrect, the task was complicated, and transcripts often seemed plausibly to match more than one target. A confounding factor in these studies is that some target locations have similarities that seem alike at some level of perception. For example, a radio telescope at the top of a hill, the observation deck of a tower, and a jetty on the edge of a bay all match a transcript description of "looking out over a long distance." A lake, a fountain, and a creek may all result in an image of water for the subject. Therefore, in several cases, even correct images may not help in the conservative differential matching procedure used.

According to the judge, the most successful procedure was a careful element-by-element comparison that tested each transcript against every target and used the transcript descriptions and drawings as arguments for or against assigning the transcript to a particular target. In most cases, this resulted in either a clear conclusion or at least a ranking of probable matches; these matches were subjected to the statistical analyses presented in this paper.

2) Summary: In summary, we do not yet have an understanding of the nature of the information-bearing signal that a subject perceives during remote viewing. The subjects commonly report that they perceive the signal visually as though they were looking at the object or place from a position in its immediate neighborhood. Furthermore, the subjects' perceptual viewpoint has mobility in that they can shift their point of view so as to describe elements of a scene that would not be visible to an observer merely standing at ground level and describing what he sees. (In particular, a subject often correctly describes elements not visible to the target demarcation team.) Finally, motion is seldom reported; in fact, moving objects often are unseen even when nearby static objects are correctly identified.

A comparison of the results of remote viewing (a so-called free-response task) with results of forced-choice tasks, such as the selection of one of four choices generated by a random number generator [58], reveals the following findings. From a statistical viewpoint, a subject is more likely to describe, with sufficient accuracy to permit blind matching, a remote site chosen at random than he is to select correctly one of four random numbers. Our experience with these phenomena leads us to consider that this difference in task performance may stem from fundamental signal-to-noise considerations. Two principal sources of noise in the system apparently are memory and imagination, both of which can give rise to mental pictures of greater clarity than the target to be perceived. In the random number task, a subject can create a perfect mental picture of each of the four possible outputs in his own imagination and then attempt to obtain the correct answer by a mental matching operation. The same is true for card guessing experiments. On the other hand, the subject in remote viewing is apparently more likely to approach the task with a blank mind as he attempts to perceive pictorial information from remote locations about which he may have no stored mental data.

Finally, we observe that most of the correct information that subjects relate to us is of a nonanalytic nature pertaining to shape, form, color, and material rather than to function or name. In consultation with Dr. Robert Ornstein of the Langley-Porter Neuropsychiatric Institute, San Francisco, CA, and with Dr. Ralph Kiernan of the Department of Neurology, Stanford University Medical Center, Stanford, CA, we have formed the tentative hypothesis that paranormal functioning may involve specialization characteristic of the brain's right hemisphere. This possibility is derived from a variety of evidence from clinical and neurosurgical sources which indicate that the two hemispheres of the human brain are specialized for different cognitive functions. The left hemisphere is predominantly active in verbal and other analytical functioning and the right hemisphere predominates in spatial and other holistic processing [59], [60]. Further research is necessary to elucidate the relationship between right hemisphere function and paranormal abilities. Nonetheless, we can say at this point that the remote-viewing results of the group of subjects at SRI have characteristics in common with more familiar performances that require right hemispheric function. The similarities include the highly schematicized drawings of objects in a room or of remote scenes. Verbal identification of these drawings is often highly inaccurate and the drawings themselves are frequently left-right reversed relative to the target configuration. Further, written material generally is not cognized. These characteristics have been seen in left brain-injured patients and in callosal-sectioned patients.

As a result of the above considerations, we have learned to urge our subjects simply to describe what they see as opposed to what they think they are looking at. We have learned that their unanalyzed perceptions are almost always a better guide to the true target than their interpretations of the perceived data.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

IV. Considerations Concerning Time

If the authors may be forgiven a personal note, we wish to express that this section deals with observations that we have been reluctant to publish because of their striking apparent incompatibility with existing concepts. The motivating factor for presenting the data at this time is the ethical consideration that theorists endeavoring to develop models for paranormal functioning should be apprised of all the observable data if their efforts to arrive at a comprehensive and correct description are to be successful.

During the course of the experimentation in remote viewing (Section III), subjects occasionally volunteered the information that they had been thinking about their forthcoming participation in a remote-viewing experiment and had an image come to them as to what the target location was to be. On these occasions, the information was given only to the experimenter remaining at SRI with the subject and was unknown to the outbound experimenter until completion of the experiment. Two of these contributions were among the most accurate descriptions turned in during those experiments. Since the target location had not yet been selected when the subject communicated his perceptions about the target, we found the data difficult to contend with.

We offer these spontaneous occurrences not as proof of precognitive perception, but rather as the motivation that led us to do further work in this field. On the basis of this firsthand evidence, together with the copious literature describing years of precognition experiments carried out in various other laboratories, we decided to determine whether a subject could perform a perceptual task that required both spatial and temporal remote viewing.

It is well known and recently has been widely discussed that nothing in the fundamental laws of physics forbids the apparent transmission of information from the future to the present (discussed further in Section V). Furthermore, there is a general dictum that "in physical law, everything that is not forbidden, is required" [61]. With this in mind, we set out to conduct very well-controlled experiments to determine whether we could deliberately design and execute experiments for the sole purpose of observing precognition under laboratory conditions.

The experimental protocol was identical to that followed in previous remote-viewing experiments with but one exception. The exception was that the subject was required to describe the remote location during a 15-min period beginning 20 min before the target was selected and 35 min before the outbound experimenter was to arrive at the target location.

In detail, as shown in Table IX, each day at ten o'clock one of the experimenters would leave SRI with a stack of ten sealed envelopes from a larger pool and randomized daily, containing traveling instructions that had been prepared, but that were unknown to the two experimenters remaining with the subject. The subject for this experiment was Hella Hammid (S4) who participated in the nine-experiment series replicating the original Price work described earlier. The traveling experimenter was to drive continuously from 10:00 until 10:30 before selecting his destination with a random number generator. (The motivation for continuous motion was our observation that objects and persons in rapid motion are not generally seen in the remote-viewing mode of perception, and we wished the traveler to be a poor target until he reached his target site.) At the end of 30 min of driving, the traveling experimenter generated a random digit from 0 to 9 with a Texas Instruments SR-51 random number generator; while still in motion, he counted down that number of envelopes and proceeded directly to the target location so as to arrive there by 10:45. He remained at the target site until 11:00, at which time he returned to the laboratory, showed his chosen target name to a security guard, and entered the experimental room.

TABLE IX
Experimental Protocol: Precognitive Remote Viewing

Time Schedule	Experimenter/Subject Activity
10:00	Outbound experimenter leaves with 10 envelopes (containing target locations) and random number generator; begins half-hour drive
10:10	Experimenters remaining with subject in the laboratory elicit from subject a description of where outbound experimenter will be from 10:45-11:00
10:25	Subject response completed, at which time laboratory part of experiment is over
10:30	Outbound experimenter obtains random number from a random number generator, counts down to associated envelope, and proceeds to target location indicated
10:45	Outbound experimenter remains at target location for 15 minutes (10:45-11:00)

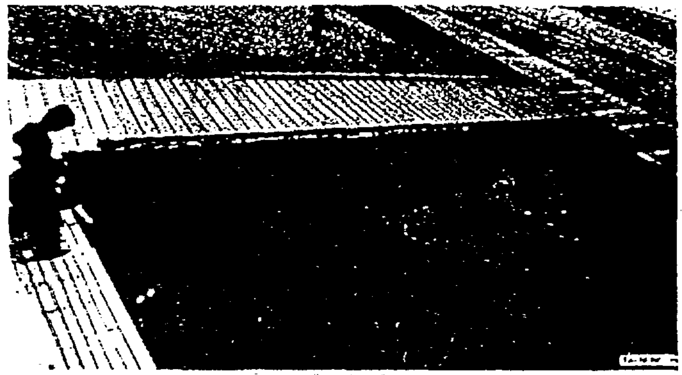


Fig. 14. Subject Hammid (S4) described "some kind of congealing tar, or maybe an area of condensed lava . . . that has oozed out to fill up some kind of boundaries."

During the same period, the protocol in the laboratory was as follows. At 10:10, the subject was asked to begin a description of the place to which the experimenter would go 35 min hence. The subject then generated a tape-recorded description and associated drawings from 10:10 to 10:25, at which time her part in the experiment was ended. Her description was thus entirely concluded 5 min before the beginning of the target selection procedure.

Four such experiments were carried out. Each of them appeared to be successful, an evaluation later verified in blind judging without error by three judges. We will briefly summarize the four experiments below.

The first target, the Palo Alto Yacht Harbor, consisted entirely of mud flats because of an extremely low tide (see Fig. 14). Appropriately, the entire transcript of the subject pertained to "some kind of congealing tar, or maybe an area of condensed lava. It looks like the whole area is covered with some kind of wrinkled elephant skin that has oozed out to fill up some kind of boundaries where (the outbound experimenter) is standing." Because of the lack of water, the dock where the remote experimenter was standing was in fact resting directly on the mud.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

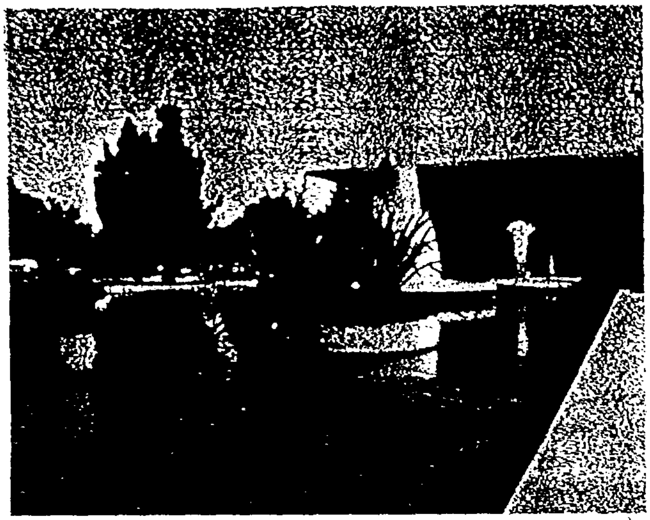


Fig. 16. Subject (S4) saw a "black iron triangle that Hal had somehow walked into" and heard a "squeak, squeak, about once a second."

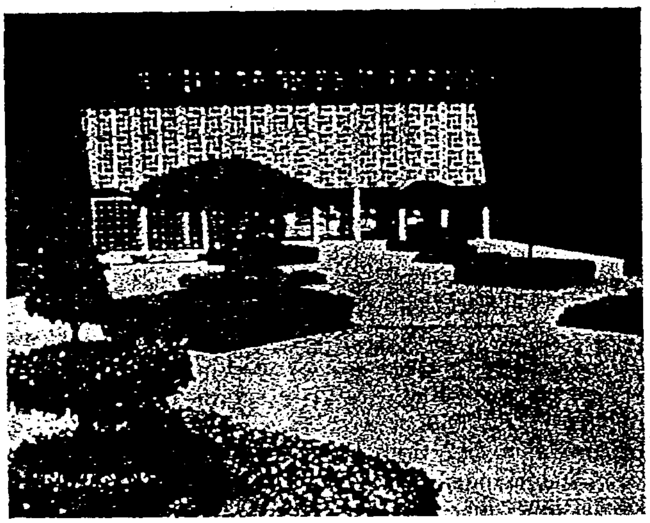


Fig. 15. Subject (S4) described a formal garden "very well manicured" behind a double colonnade.

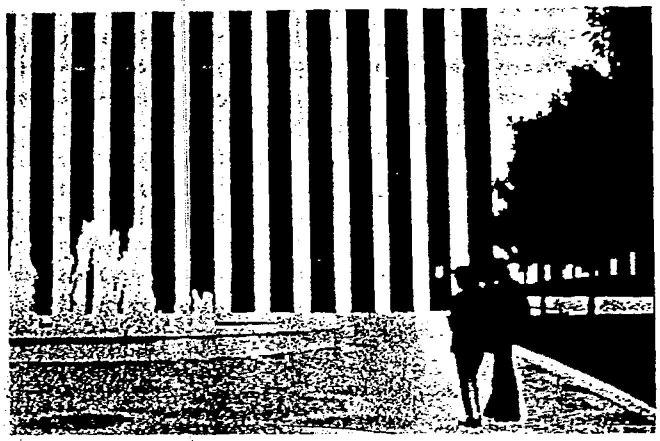


Fig. 17. Subject (S4) described a very tall structure located among city streets and covered with "Tiffany-like glass."

Note that the subject has learned not to rush into interpretation as to the nature or purpose of the place. This is a result of our cautioning based on the observation that such efforts tend to be purely analytical and in our experience are almost invariably incorrect. If a subject can limit himself to what he sees, he is often then able to describe a scene with sufficient accuracy that an observer can perform the analysis for him and identify the place.

The second target visited was the fountain at one end of a large formal garden at Stanford University Hospital (Fig. 15). The subject gave a lengthy description of a formal garden behind a wall with a "double colonnade" and "very well manicured." When we later took the subject to the location, she was herself taken aback to find the double colonnaded wall leading into the garden just as described.

The third target was a children's swing at a small park 4.6 km from the laboratory (Fig. 16). The subject repeated again and again that the main focus of attention at the site was a "black iron triangle that the outbound experimenter had somehow walked into or was standing on." The triangle was "bigger than a man," and she heard a "squeak, squeak, about once a second," which we observe is a match to the black metal swing that did squeak.

The final target was the Palo Alto City Hall (Fig. 17). The subject described a very, very tall structure covered with "Tiffany-like glass." She had it located among city streets and with little cubes at the base. The building is glass-covered, and the little cubes are a good match to the small elevator exit buildings located in the plaza in front of the building.

To obtain a numerical evaluation of the accuracy of the precognitive viewing, the experimental results were subjected to independent judging on a blind basis by three SRI scientists who were not otherwise associated with the experiment. The judges were asked to match the four locations, which they visited, against the unedited typed manuscripts of the tape-recorded narratives, along with the drawings generated by the remote viewer. The transcripts were presented unlabeled and in random order and were to be used without replacement. A correct match required that the transcript of a given experiment be matched with the target of that experiment. All three judges independently matched the target data to the response data without error. Under the null hypothesis (no information channel and a random selection of descriptions without replacement), each judge independently obtained a result significant at $p = (4!)^{-1} = 0.042$.

For reasons we do not as yet understand, the four transcripts generated in the precognition experiment show exceptional coherence and accuracy as evidenced by the fact that all of the judges were able to match successfully all of the transcripts to

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the corresponding target locations. A long-range experimental program devoted to the clarification of these issues and involving a number of subjects is under way. The above four experiments are the first four carried out under this program.

Currently, we have no precise model of this spatial and temporal remote-viewing phenomenon. However, models of the universe involving higher order synchronicity or correlation have been proposed by the physicist Pauli and the psychologist Carl Jung [62].

> ACAUSALITY. If natural law[5] were an absolute truth, then of course there could not possibly be any processes that deviate from it. But since causality[5] is a *statistical* truth, it holds good only on average and thus leaves room for *exceptions* which must somehow be experienceable, that is to say, *real*. I try to regard synchronistic events as acausal exceptions of this kind. They prove to be relatively independent of space and time; they relativize space and time insofar as space presents in principle no obstacle to their passage and the sequence of events in time is inverted so that it looks as if an event which has not yet occurred were causing a perception in the present.

We shall see in the next section that such a description, though poetic, has some basis in modern physical theory.

V. Discussion

It is important to note at the outset that many contemporary physicists are of the view that the phenomena that we have been discussing are not at all inconsistent with the framework of physics as currently understood. In this emerging view, the often-held belief that observations of this type are incompatible with known laws *in principle* is erroneous, such a concept being based on the naive realism prevalent before the development of modern quantum theory and information theory.

One hypothesis, put forward by I. M. Kogan of the USSR, is that information transfer under conditions of sensory shielding is mediated by extremely low-frequency (ELF) electromagnetic waves in the 300–1000-km region [37]-[40]. Experimental support for the hypothesis is claimed on the basis of slower than inverse square attenuation, compatible with source-percipient distances lying in the induction field range as opposed to the radiation field range; observed low bit rates (0.005–0.1 bit/s) compatible with the information carrying capacity of ELF waves; apparent ineffectiveness of ordinary electromagnetic shielding as an attenuator; and standard antenna calculations entailing biologically generated currents yielding results compatible with observed signal-to-noise ratios.

M. Persinger, Psychophysiology Laboratory, Laurentian University, Toronto, Canada, has narrowed the ELF hypothesis to the suggestion that the 7.8-Hz "Shumann waves" and their harmonics propagating along the earth–ionosphere waveguide duct may be responsible. Such an hypothesis is compatible with driving by brain-wave currents and leads to certain other hypotheses such as asymmetry between east-west and west-east propagation, preferred experimental times (midnight–4 A.M.), and expected negative correlation between success and the U index (a measure of geomagnetic disturbance throughout the world). Persinger claims initial support for these factors on the basis of a literature search [63], [64].

On the negative side with regard to a straightforward ELF interpretation as a blanket hypothesis are the following: a) apparent real-time descriptions of remote activities in sufficient detail to require a channel capacity in all probability greater than that allowed by a conventional modulation of an ELF signal; b) lack of a proposed mechanism for coding and decoding the information onto the proposed ELF carrier; and c) apparent precognition data. The hypothesis must nonetheless remain open at this stage of research, since it is conceivable that counterindication a) may eventually be circumvented on the basis that the apparent high bit rate results from a mixture of low bit rate input and high bit rate "filling in the blanks" from imagination; counterindication b) is common to a number of normal perceptual tasks and may therefore simply reflect a lack of sophistication on our part with regard to perceptual functioning [65]; and counterindication c) may be accommodated by an ELF hypothesis if advanced waves as well as retarded waves are admitted [66], [67]. Experimentation to determine whether the ELF hypothesis is viable can be carried out by the use of ELF sources as targets, by the study of parametric dependence on propagational directions and diurnal timing, and by the exploration of interference effects caused by creation of a high-intensity ELF environment during experimentation, all of which are under consideration in our laboratory and elsewhere.

Some physicists believe that the reconciliation of observed paranormal functioning with modern theory may take place at a more fundamental level—namely, at the level of the foundations of quantum theory. There is a continuing dialog, for example, on the proper interpretation of the effect of an observer (consciousness) on experimental measurement [68], and there is considerable current interest in the implications for our notions of ordering in time and space brought on by the observation [69], [70] of nonlocal correlation or "quantum interconnectedness" (to use Bohm's term [71]) of distant parts of quantum systems of macroscopic dimensions. The latter, Bell's theorem [72], emphasizes that "no theory of reality compatible with quantum theory can require spatially separated events to be independent" [73], but must permit interconnectedness of distant events in a manner that is contrary to ordinary experience [74]-[75]. This prediction has been experimentally tested and confirmed in the recent experiments of, for example, Freedman and Clauser [69], [70].

E. H. Walker and O. Costa de Beauregard, independently proposing theories of paranormal functioning based on quantum concepts, argue that observer effects open the door to the possibility of nontrivial coupling between consciousness and the environment and that the nonlocality principle permits such coupling to transcend spatial and temporal barriers [76], [77].

Apparent "time reversibility"—that is, effects (e.g., observations) apparently preceding causes (e.g., events)—though conceptually difficult at first glance, may be the easiest of apparent paranormal phenomena to assimilate within the current theoretical structure of our world view. In addition to the familiar retarded potential solutions $f(t - r/c)$, it is well known that the equations of, for example, the electromagnetic field admit of advanced potential solutions $f(t + r/c)$—solutions that would appear to imply a reversal of cause and effect. Such solutions are conventionally discarded as not corresponding to any observable physical event. One is cautioned, however, by statements such as that of Stratton in his basic text on electromagnetic theory [78].

[5] As usually understood.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The reader has doubtless noted that the choice of the function $f(t - r/c)$ is highly arbitrary, since the field equation admits also a solution $f(t + r/c)$. This function leads obviously to an advanced time, implying that the field can be observed before it has been generated by the source. The familiar chain of cause and effect is thus reversed and this alternative solution might be discarded as logically inconceivable. However, the application of "logical" causality principles offers very insecure footing in matters such as these and we shall do better to restrict the theory to retarded action solely on the grounds that this solution alone conforms to the present physical data.

Such caution is justified by the example in the early 1920's of Dirac's development of the mathematical description of the relativistic electron that also yielded a pair of solutions, one of which was discarded as inapplicable until the discovery of the positron in 1932.

In an analysis by O. Costa de Beauregard, an argument is put forward that advanced potentials constitute a convergence toward "finality" in a manner symmetrical to the divergence of retarded potentials as a result of causality [77]. Such phenomena are generally unobservable, however, on the gross macroscopic scale for statistical reasons. This is codified in the thermodynamic concept that for an isolated system entropy (disorder) on the average increases. It is just this requirement of isolation, however, that has been weakened by the observer problem in quantum theory, and O. Costa de Beauregard argues that the finality principle is maximally operative in just those situations where the intrusion of consciousness as an ordering phenomenon results in a significant local reversal of entropy increase. At this point, further discussion of the subtleties of such considerations, though apropos, would take us far afield, so we simply note that such advanced waves, if detected, could in certain cases constitute a carrier of information precognitive to the event.

The above arguments are not intended to indicate that the precise nature of the information channel coupling remote events and human perception is understood. Rather, we intend to show only that modern theory is not without resources that can be brought to bear on the problems at hand, and we expect that these problems will, with further work, continue to yield to analysis and specification.

Furthermore, independent of the mechanisms that may be involved in remote sensing, observation of the phenomenon implies the existence of an information channel in the information-theoretic sense. Since such channels are amenable to analysis on the basis of communication theory techniques, as indicated earlier, channel characteristics such as bit rate can be determined independent of a well-defined physical channel model in the sense that thermodynamic concepts can be applied to the analysis of systems independent of underlying mechanisms. Furthermore, as we have seen from the work of Ryzl discussed in Section II, it is possible to use such a channel for error-free transmission of information if redundancy coding is used. (See also Appendix A.) Therefore, experimentation involving the collection of data under specified conditions permits headway to be made despite the formidable work that needs to be done to clarify the underlying bases of the phenomena.

VI. Conclusion

For the past three years we have had a program in the Electronics and Bioengineering Laboratory of SRI to investigate those facets of human perception that appear to fall outside the range of well-understood perceptual or processing capabilities. The primary achievement of this program has been the elicitation of high-quality "remote viewing"—the ability of both experienced subjects and inexperienced volunteers to view, by means of innate mental processes, remote geographical or technical targets such as roads, buildings, and laboratory apparatus. Our accumulated data from over fifty experiments with more than a half-dozen subjects indicate the following. a) The phenomenon is not a sensitive function of distance over a range of several kilometers. b) Faraday cage shielding does not appear to degrade the quality or accuracy of perception. c) Most of the correct information that subjects relate is of a nonanalytic nature pertaining to shape, form, color, and material rather than to function or name. (This aspect suggests a hypothesis that information transmission under conditions of sensory shielding may be mediated primarily by the brain's right hemisphere.) d) The principal difference between experienced subjects and inexperienced volunteers is *not* that the latter never exhibit the faculty, but rather that their results are simply less reliable. (This observation suggests the hypothesis that remote viewing may be a latent and widely distributed, though repressed, perceptual ability.)

Although the precise nature of the information channel coupling remote events and human perception is not yet understood, certain concepts in information theory, quantum theory, and neurophysiological research appear to bear directly on the issue. As a result, the working assumption among researchers in the field is that the phenomena of interest is consistent with modern scientific thought, and can therefore be expected to yield to the scientific method. Further, it is recognized that communication theory provides powerful techniques, such as the use of redundancy coding to improve signal-to-noise ratio, which can be employed to pursue special-purpose application of the remote-sensing channel independent of an understanding of the underlying mechanisms. We therefore consider it important to continue data collection and to encourage others to do likewise; investigations such as those reported here need replication and extension under as wide a variety of rigorously controlled conditions as possible.

Appendix A
Signal Enhancement in a Paranormal Communication Channel by Application of Redundancy Coding

Independent of the mechanisms that may be involved in remote sensing, observation of the phenomenon implies the existence of an information channel in the information-theoretic sense. As we have seen from the work of Ryzl discussed in Section II,[6] it is even possible to use such a (noisy) channel for error-free transmission of information if sufficient redundancy coding is used [30], [31]. Following is a general procedure that we have used successfully for signal enhancement.

We shall assume that the "message" consists of a stream of binary digits (0,1) of equal probability (e.g., binary sort of green/white cards as in Ryzl's case, English text encoded as in Table X and sent long distance by strobe light on/off, and so on). To combat channel noise, each binary digit to be sent through the channel requires the addition of redundancy bits (coding). Efficient coding requires a compromise between the desire to maximize reliability and the desire to minimize re-

[6] See also the note added in proof on the successful work done by Carpenter.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE X
5-BIT CODE FOR ALPHANUMERIC CHARACTERS

E	00000	Y	01000
T	11111	G,J	10111
N	00001	W	01001
R	11110	V	10110
I	00010	B	01010
O	11101	Ø	10101
A	00011	1	01011
S,X,Z	11100	2	10100
D	00100	3	01100
H	11011	4	10011
L	00101	5	01101
C,K,Q	11010	6	10010
F	00110	7	01110
P	11001	8	10001
U	00111	9	01111
M	11000	.	10000

Note: Alphabet characters listed in order of decreasing frequency in English text. See, for example, A. Sinkov [79]. (The low-frequency letters, X, Z, K, Q, and J, have been grouped with similar characters to provide space for numerics in a 5-bit code.) In consideration of the uneven distribution of letter frequencies in English text, this code is chosen such that 0 and 1 have equal probability.

dundancy. One efficient coding scheme for such a channel is obtained by application of a sequential sampling procedure of the type used in production-line quality control [80]. The adaptation of such a procedure to paranormal communication channels, which we now discuss, was considered first by Taetzsch [81]. The sequential method gives a rule of procedure for making one of three possible decisions following the receipt of each bit: accept 1 as the bit being transmitted; reject 1 as the bit being transmitted (i.e., accept 0); or continue transmission of the bit under consideration. The sequential sampling procedure differs from fixed-length coding in that the number of bits required to reach a final decision on a message bit is not fixed before transmission, but depends on the results accumulated with each transmission. The principal advantage of the sequential sampling procedure as compared with the other methods is that, on the average, fewer bits per final decision are required for an equivalent degree of reliability.

Use of the sequential sampling procedure requires the specification of parameters that are determined on the basis of the following considerations. Assume that a message bit (0 or 1) is being transmitted. In the absence of *a priori* knowledge, we may assume equal probability ($p = 0.5$) for the two possibilities (0,1). Therefore, from the standpoint of the receiver, the probability of correctly identifying the bit being transmitted is $p = 0.5$ because of chance alone. An operative remote-sensing channel could then be expected to alter the probability of correct identification to a value $p = 0.5 + \psi$, where the parameter ψ satisfies $0 < |\psi| < 0.5$. (The quantity may be positive or negative depending on whether the paranormal channel results in so-called psi-hitting or psi-missing.) Good psi functioning on a repetitive task has been observed to result in $\psi = 0.12$, as reported by Ryzl [31]. Therefore, to indicate the design procedure, let us assume a baseline psi parameter $\psi_b = 0.1$ and design a communication system on this basis.

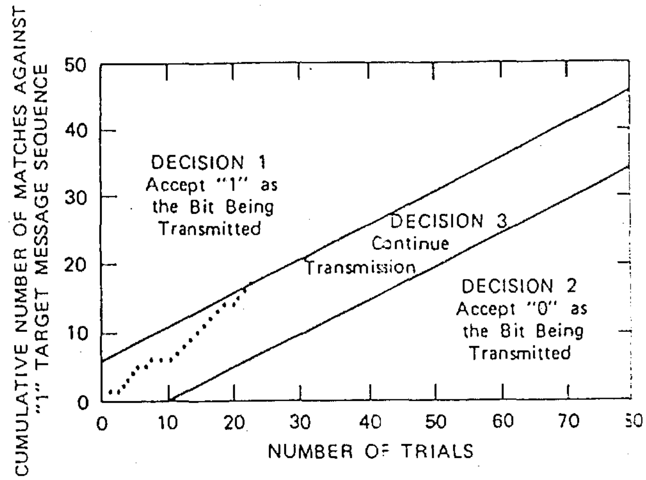


Fig. 18. Enhancement of signal-to-noise ratio by sequential sampling procedure ($p_0 = 0.4, p_1 = 0.6, \alpha = 0.01, \beta = 0.01$).

The question to be addressed is whether, after repeated transmission, a given message bit is labeled a "1" at a low rate p_0 commensurate with the hypothesis H_0 that the bit in question is a "0," or at a higher rate p_1 commensurate with the hypothesis H_1 that the bit in question is indeed a "1." The decision-making process requires the specification of four parameters.

- p_0 The probability of labeling incorrectly a "0" message bit as a "1." The probability of labeling correctly a "0" as a "0" is $p = 0.5 + \psi_b = 0.6$. Therefore, the probability of labeling incorrectly a "0" as a "1" is $1 - p = 0.4 = p_0$.
- p_1 The probability of labeling correctly a "1" message bit as a "1," is given by $p_1 = 0.5 + \psi_b = 0.6$.
- α The probability of rejecting a correct identification for a "0" (Type I error). We shall take $\alpha = 0.01$.
- β The probability of accepting an incorrect identification for a "1" (Type II error). We shall take $\beta = 0.01$.

With the parameters thus specified, the sequential sampling procedure provides for construction of a decision graph as shown in Fig. 18. The equations for the upper and lower limit lines are

$$\Sigma_1 = d_1 + SN$$

$$\Sigma_0 = -d_0 + SN$$

where

$$d_1 = \frac{\log \dfrac{1-\beta}{\alpha}}{\log \dfrac{p_1}{p_0}\dfrac{1-p_0}{1-p_1}} \qquad d_0 = \frac{\log \dfrac{1-\alpha}{\beta}}{\log \dfrac{p_1}{p_0}\dfrac{1-p_0}{1-p_1}}$$

$$S = \frac{\log \dfrac{1-p_0}{1-p_1}}{\log \dfrac{p_1}{p_0}\dfrac{1-p_0}{1-p_1}}$$

in which S is the slope, N is the number of trials, and d_1 and d_0 are the y-axis intercepts. A cumulative record of receiver-generated responses to the target bit is compiled until either

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

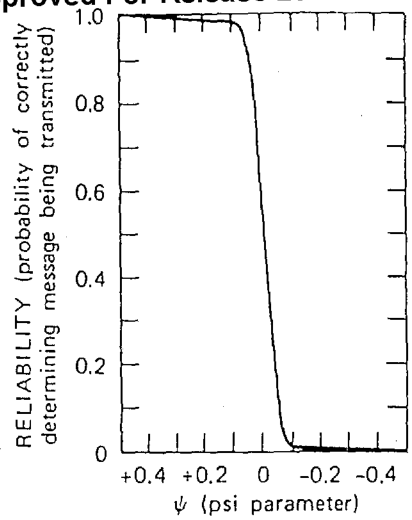


Fig. 19. Reliability curve for sequential sampling procedure ($p_0 = 0.4$, $p_1 = 0.6$, $\alpha = 0.01$, $\beta = 0.01$).

the upper or the lower limit line is reached, at which point a decision is made to accept 0 or 1 as the bit being transmitted.

Channel reliability (probability of correctly determining message being transmitted) as a function of operative psi parameter ψ is plotted in Fig. 19. As observed, the sequential sampling procedure can result in 90 percent or greater reliability with psi parameters on the order of a few percent.

Implementation of the sequential sampling procedure requires the transmission of a message coded in binary digits. Therefore, the target space must consist of dichotomous elements such as the white and green cards used in the experiments by Ryzl.

In operation, a sequence corresponding to the target bit (0 or 1) is sent and the cumulative entries are made (Fig. 18) until a decision is reached to accept either a 1 or a 0 as the bit being transmitted. At a prearranged time, the next sequence is begun and continues as above until the entire message has been received. A useful alternative, which relieves the percipient of the burden of being aware of his self-contradiction from trial to trial, consists of cycling through the entire message repetitively and entering each response on its associated graph until a decision has been reached on all message bits. The authors have used this technique successfully in a pilot study, but a discussion of this would take us beyond the intended scope of this paper.

From the results obtained in such experiments, the channel bit rate can be ascertained for the system configuration under consideration. Furthermore, bit rates for other degrees of reliability (i.e., for other p_0, p_1, α, and β) can be estimated by construction of other decision curves over the same data base and thus provide a measure of the bit rate per degree of reliability.

In summary, the procedures described here can provide for a specification of the characteristics of a remote-sensing channel under well-defined conditions. These procedures also provide for a determination of the feasibility of such a channel for particular applications.

Appendix B

Remote-Viewing Transcript

Following is the unedited transcript of the *first* experiment with an SRI volunteer (S6), a mathematician in the computer science laboratory, with no previous experience in remote viewing. The target, determined by random procedure, was White's Plaza, a plaza with fountain at Stanford University (shown in Fig. 8). As is our standard protocol, the experimenter with the subject is kept ignorant of the specific target visited as well as the contents of the target pool. The experimenter's statements and questions are italics.

Today is Monday, October 7th. It is 11:00 and this is a remote viewing experiment with Russ Targ, Phyllis Cole, and Hal Puthoff. In this experiment Hal will drive to a remote site chosen by a random process. Phyllis Cole will be the remote viewer, and Russ Targ is the monitor. We expect this experiment to start at twenty minutes after eleven and run for fifteen minutes.

It is just about twenty minutes after eleven and Hal should be at his target location by now.

Why don't you tell me what kind of pictures you see and what you think he might be doing or experiencing.

The first thing that came to mind was some sort of a large, square kind of a shape. Like Hal was in front of it. It was a . . . not a building or something, it was a square. I don't know if it was a window, but something like that so that the bottom line of it was not at the ground. About where his waist was, at least. That's what it seemed to me. It seems outdoors somehow. Tree.

Does Hal seem to be looking at that square?

I don't know. The first impression was that he wasn't, but I have a sense that whatever it was was something one might look at. I don't know if it would be a sign, but something that one might look at.

Can you tell if it is on the ground or vertical?

It seemed vertical.

I don't have a sense that it was part of anything particular. It might be on a building or part of a building, but I don't know. There was a tree outside, but I also got the impression of cement. I don't have the impression of very many people or traffic either. I have the sense that he is sort of walking back and forth. I don't have any more explicit picture than that.

Can you move into where he is standing and try to see what he is looking at?

I picked up he was touching something—something rough. Maybe warm and rough. Something possibly like cement.

It is twenty-four minutes after eleven.

Can you change your point of view and move above the scene so you can get a bigger picture of what's there?

I still see some trees and some sort of pavement or something like that. Might be a courtyard. The thing that came to mind was it might be one of the plazas at Stanford campus or something like that, cement.

Some kinds of landscaping.

I said Stanford campus when I started to see some things in White Plaza, but I think that is misleading.

I have the sense that he's not moving around too much. That it's in a small area.

I guess I'll go ahead and say it, but I'm afraid I'm just putting on my impressions from Stanford campus. I had the impression of a fountain. There are two in the plaza, and it seemed that Hal was possibly near the, what they call Mem Claw.

What is that?

It's a fountain that looks rather like a claw. It's a black sculpture. And it has benches around it made of cement.

Are there any buildings at the place you are looking at? Are there any buildings? You described a kind of a courtyard.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Usually at some places there should be a building, large or small that the courtyard is about. Look at the end or the sides of the courtyard. Is there anything to be seen?

I have a sense that there are buildings. It's not solid buildings. I mean there are some around the periphery and I have a sense that none of them are very tall. Maybe mostly one story, maybe an occasional two story one.

Do you have any better idea of what your square was that you saw at the outset?

No. I could hazard different kinds of guesses.

Does it seem part of this scene?

It . . . I think it could be. It could almost be a bulletin board or something with notices on it maybe.

Or something that people are expected to look at. Maybe a window with things in it that people were expected to look at.

What kind of trees do you see in this place?

I don't know what kind they are. The impression was that they were shade trees and not terribly big. Maybe 12 feet of trunk and then a certain amount of branches above that. So that the branches have maybe a 12 foot diameter, or something. Not real big trees.

New trees rather than old trees?

Yeah, maybe 5 or 10 years old, but not real old ones.

Is there anything interesting about the pavement?

No. It seems to be not terribly new or terribly old. Not very interesting. There seems to be some bits of landscaping around. Little patches of grass around the edges and peripheries. Maybe some flowers. But, not lush.

You saw some benches. Do you want to tell me about them?

Well, that's my unsure feeling about this fountain. There was some kind of benches of cement. Curved benches, it felt like.

They were of rough cement.

What do you think Hal is doing while he is there?

I have a sense that he is looking at things trying to project them. Looking at different things and sort of walking back and forth not covering a whole lot of territory.

Sometimes standing still while he looks around.

I just had the impression of him talking, and I almost sense that it was being recorded or something. I don't know if he has a tape recorder, but if it's not that, then he is saying something because it needed to be remembered. *It's 11:33. He's just probably getting ready to come back.*

ACKNOWLEDGMENT

The authors wish to thank the principal subjects, Mrs. Hella Hammid, Pat Price, and Ingo Swann, who showed patience and forbearance in addition to their enthusiasm and outstanding perceptual abilities. We note with sadness the death of one of our subjects, Mr. Price. We express our sincere thanks also to Earle Jones, Bonnar Cox, and Dr. Arthur Hastings, of SRI, and Mrs. Judith Skutch and Richard Bach, without whose encouragement and support this work could not have taken place.

REFERENCES

[1] J. R. Smythies, Ed., *Science and ESP.* London, England: Routledge, 1967.

[2] C. Evans, "Parapsychology—What the questionnaire revealed," *New Scientist*, Jan. 25, 1973, p. 209.

[3] A. Gauld, *The Founders of Psychical Research.* New York: Schocken Books, 1968. See also W. Crookes, *Researches in the Phenomena of Spiritualism.* London, England: J. Burns, 1874.

[4] R. Targ and H. Puthoff, "Information transmission under conditions of sensory shielding," *Nature*, vol. 252, pp. 602-607, Oct. 18, 1974.

[5] D. D. Home, *Lights and Shadows of Spiritualism.* New York: G. W. Carleton, 1877.

[6] J. Coover, *Experiments in Psychical Research.* Palo Alto, CA: Stanford Univ. Press, 1917.

[7] G. Estabrooks, *Bull. Boston Society for Psychical Research*, 1927. See also [12, pp. 18-19].

[8] L. T. Troland, *Techniques for the Experimental Study of Telepathy and Other Alleged Clairvoyant Processes.* Albany, NY, 1928.

[9] J. B. Rhine, *New Frontiers of the Mind.* New York: Farrar and Rinehart, 1937.

[10] J. Pratt and J. B. Rhine *et al.*, *Extra-Sensory Perception after Sixty Years.* New York: Henry Holt, 1940.

[11] C. Scott, "G. Spencer Brown and probability: A critique," *J. Soc. Psychical Res.*, vol. 39, pp. 217-234, 1958.

[12] G. R. Price, "Science and the supernatural," *Science*, vol. 122, pp. 359-367, 1955.

[13] ——, "Apology to Rhine and Soal," *Science*, vol. 175, p. 359, 1972.

[14] J. B. Rhine, "A new case of experimenter unreliability," *J. Parapsychol.*, vol. 38, pp. 215-225, June 1974.

[15] S. G. Soal and F. Bateman, *Modern Experiments in Telepathy.* London, England: Faber and Faber, 1953.

[16] C. Scott and P. Haskell, "'Normal' explanation of the Soal-Goldney experiments in extra-sensory perception," *Nature*, vol. 245, pp. 52-54, Sept. 7, 1973.

[17] C. E. M. Hansel, *ESP—A Scientific Evaluation.* New York: Scribner, 1966.

[18] J. B. Rhine and J. G. Pratt, "A review of the Pearce-Pratt distance series of ESP tests," *J. Parapsychol.*, vol. 18, pp. 165-177, 1954.

[19] J. G. Pratt and J. L. Woodruff, "Size of stimulus symbols in extrasensory perception," *J. Parapsychol.*, vol. 3, pp. 121-158, 1939.

[20] S. G. Soal and H. T. Bowden, *The Mind Readers: Recent Experiments in Telepathy.* New Haven, CT: Yale Univ. Press, 1954.

[21] C. Honorton, "Error some place!" *J. Commun.*, vol. 25, no. 1 (Annenberg School of Commun.), Winter 1975.

[22] M. Ryzl, "Training the psi faculty by hypnosis," *J. Amer. Soc. Psychical Res.*, vol. 41, pp. 234-251, 1962.

[23] *CIBA Foundation Symposium on Extra Sensory Perception.* Boston, MA: Little, Brown, 1956.

[24] M. Ryzl and J. Pratt, "A repeated-calling ESP test with sealed cards," *J. Parapsychol.*, vol. 27, pp. 161-174, 1963.

[25] ——, "A further confirmation of stabilized ESP performance in a selected subject," *J. Parapsychol.*, vol. 27, pp. 73-83, 1963.

[26] J. Pratt, "Preliminary experiments with a 'borrowed' ESP subject," *J. Amer. Soc. Psychical Res.*, vol. 42, pp. 333-345, 1964.

[27] J. Pratt and J. Blom, "A confirmatory experiment with 'borrowed' outstanding ESP subject," *J. Amer. Soc. Psychical Res.*, vol. 42, pp. 381-388, 1964.

[28] W. G. Roll and J. G. Pratt, "An ESP test with aluminum targets," *J. Amer. Soc. Psychical Res.*, vol. 62, pp. 381-387, 1968.

[29] J. Pratt, "A decade of research with a selected ESP subject: An overview and reappraisal of the work with Pavel Stepanek," *Proc. Amer. Soc. Psychical Res.*, vol. 30, 1973.

[30] C. Shannon and W. Weaver, *The Mathematical Theory of Communication.* Urbana, IL: Univ. Illinois Press, 1949.

[31] M. Ryzl, "A model for parapsychological communication," *J. Parapsychol.*, vol. 30, pp. 18-31, Mar. 1966.

[32] C. Tart, "Card guessing tests: Learning paradigm or extinction paradigm," *J. Amer. Soc. Psychical Res.*, vol. 60, p. 46, 1966.

[33] M. Ulman and S. Krippner, with A. Vaughan, *Dream Telepathy.* New York: Macmillan, 1973.

[34] C. Honorton, "State of awareness factors in psi activation," *J. Amer. Soc. Psychical Res.*, vol. 68, pp. 246-257, 1974.

[35] *Proc. 2nd Int. Congr. Psychotronic Research* (Monte Carlo). Cotati, CA: Int. Assoc. Psychotronic Res., 1975.

[36] L. L. Vasiliev, *Experiments in Mental Suggestion.* Hampshire, England: ISMI Publ., 1963.

[37] I. M. Kogan, "Is telepathy possible?" *Radio Eng.*, vol. 21, p. 75, Jan. 1966.

[38] ——, "Telepathy, hypotheses and observations," *Radio Eng.*, vol. 22, p. 141, Jan. 1967.

[39] ——, "Information theory analysis of telepathic communication experiments," *Radio Eng.*, vol. 23, p. 122, Mar. 1968.

[40] ——, "The information theory aspect of telepathy," RAND Publ., Santa Monica, CA, p. 4145, July 1969.

[41] A. S. Presman, *Electromagnetic Fields and Life.* New York: Plenum, 1970.

[42] Y. A. Kholodov, Ed., *Influence of Magnetic Fields on Biological Objects*, JPRS 63038, NTIS, Springfield, VA, Sept. 24, 1974.

[43] Y. A. Kholodov, "Investigation of the direct effect of magnetic fields on the central nervous system," in *Proc. 1st Conf. Psychotronic Res.*, JPRS L/5022-1 and 2, Sept. 6, 1974.

[44] D. Mennie, "Consumer electronics," *IEEE Spectrum*, vol. 12, pp. 34-35, Mar. 1975.

[45] W. P. Zinchenko, A. N. Leontiev, B. M. Lomov, and A. R. Luria,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

"Par[illegible]" [illegible] vol. 9, pp. 128–136, 1973.

[46] R. Cavanna, Ed., *Proc. Int. Conf. Methodology in PSI Research.* New York: Parapsychology Foundation, 1970.

[47] E. D. Dean, "Plethysmograph recordings as ESP responses," *Int. J. Neuropsychiatry*, vol. 2, Sept. 1966.

[48] C. Tart, "Physiological correlates of psi cognition," *Int. J. Parapsychol.*, no. 4, 1963.

[49] D. H. Lloyd, "Objective events in the brain correlating with psychic phenomena," *New Horizons*, vol. 1, no. 2, Summer 1973.

[50] J. Silverman and M. S. Buchsbaum, "Perceptual correlates of consciousness; A conceptual model and its technical implications for psi research," in *Psi Favorable States of Consciousness*, R. Cavanna, Ed. New York: Parapsychology Foundation, pp. 143–169, 1970.

[51] J. Kamiya, "Comment to Silverman and Buchsbaum," *ibid.*, pp. 158–159.

[52] D. Hill and G. Parr, *Electroencephalography, A Symposium on Its Various Aspects.* New York: Macmillan, 1963.

[53] T. D. Duane and T. Behrendt, "Extrasensory electroencephalographic induction between identical twins," *Science*, vol. 150, p. 367, 1965.

[54] K. Osis, *ASPR Newsletter*, no. 14, 1972.

[55] R. L. Morris, "An exact method for evaluating preferentially matched free-response material," *J. Amer. Soc. Psychical Res.*, vol. 66, p. 401, Oct. 1972.

[56] G. R. Schmeidler, "PK effects upon continuously recorded temperatures," *J. Amer. Soc. Psychical Res.*, vol. 67, no. 4, Oct. 1973.

[57] W. Scherer, "Spontaneity as a factor in ESP," *J. Amer. Soc. Psychical Res.*, vol. 12, pp. 126–147, 1948.

[58] R. Targ, P. Cole, and H. Puthoff, "Techniques to enhance man/machine communication," SRI, Menlo Park, CA, Final Rep., NASA Contract NAS7-100, June 1974.

[59] R. Ornstein, *The Nature of Human Consciousness.* San Francisco, CA: Freeman, 1973, ch. 7 and 8.

[60] R. W. Sperry, "Cerebral organization and behavior," *Science*, vol. 133, pp. 1749–1757, 1961.

[61] O. Bilaniuk and E. C. G. Sudarshan, "Particles beyond the light barrier," *Phys. Today*, vol. 22, May 5, 1969.

[62] W. Pauli and C. G. Jung, Eds., *The Interpretation of Nature and the Psyche* (Bollingen Ser. LI). Princeton, NJ: Princeton Univ. Press, 1955.

[63] M. A. Persinger, "ELF waves and ESP," *New Horizons Trans. Toronto Society for Psychical Research*, vol. 1, no. 5, Jan. 1975.

[illegible] Mechanisms and models," M.S.S. Information Corp., New York, 1974.

[65] B. Julesz, *Foundations of Cyclopean Perception.* Chicago, IL: Univ. Chicago Press, 1971.

[66] H. Puthoff and R. Targ, in *Psychic Exploration—A Challenge for Science*, J. White, Ed. New York: Putnam, 1974, pp. 522–542.

[67] G. Feinberg, "Precognition—A memory of things future?" in *Proc. Conf. Quantum Physics and Parapsychology* (Geneva, Switzerland). New York: Parapsychology Foundation, 1975.

[68] E. P. Wigner, "The problem of measurement," *Amer. J. Phys.*, vol. 31, no. 1, p. 6, 1963.

[69] J. J. Freedman and J. F. Clauser, "Experimental test of local hidden variable theories," *Phys. Rev. Lett.*, vol. 28, no. 14, p. 938, Apr. 3, 1972.

[70] J. F. Clauser and M. A. Horne, "Experimental consequences of objective local theories," *Phys. Rev. D*, vol. 10, no. 2, p. 526, July 15, 1974.

[71] D. Bohm and B. Hiley, "On the intuitive understanding of nonlocality as implied by quantum theory" (Birkbeck College, London, England), Feb. 1974, Preprint.

[72] J. S. Bell, "On the problem of hidden variables in quantum theory," *Rev. Mod. Phys.*, vol. 38, no. 3, p. 447, July 1966.

[73] H. Stapp, "Theory of reality," Lawrence-Berkeley Lab. Rep. LBL-3837, Univ. California, Berkeley, Apr. 1975.

[74] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Phys. Rev.*, vol. 47, p. 777, May 15, 1935.

[75] R. H. Dicke and J. P. Wittke, *Introduction to Quantum Mechanics.* Reading, MA: Addison-Wesley, 1960, ch. 7.

[76] E. H. Walker, "Foundations of paraphysical and parapsychological phenomena," in *Proc. Conf. Quantum Physics and Parapsychology* (Geneva, Switzerland). New York: Parapsychology Foundation, 1975.

[77] O. Costa de Beauregard, "Time symmetry and interpretation of quantum mechanics," Lecture delivered at Boston Colloquim for Philosophy of Science (Feb. 1974), *Foundations of Physics* (in press).

[78] J. A. Stratton, *Electromagnetic Theory.* New York: McGraw-Hill, 1941.

[79] A. Sinkov, *Elementary Cryptanalysis—A Mathematical Approach.* New York: Random House, 1968.

[80] P. Hoel, *Introduction to Mathematical Statistics*, 2nd ed. New York: Wiley, 1954, p. 27.

[81] R. Taetzsch, "Design of a psi communications system," *Int. J. Parapsychol.*, vol. 4, no. 1, p. 35, Winter 1962.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The Persistent Paradox of Psychic Phenomena: An Engineering Perspective

ROBERT G. JAHN

Invited Paper

Abstract—Although a variety of so-called psychic phenomena have attracted man's attention throughout recorded history, organized scholarly effort to comprehend such effects is just one century old, and systematic academic research roughly half that age. Over recent years, a sizeable spectrum of evidence has been brought forth from reputable laboratories in several disciplines to suggest that at times human consciousness can acquire information inaccessible by any known physical mechanism (ESP), and can influence the behavior of physical systems or processes (PK), but even the most rigorous and sophisticated of these studies display a characteristic dilemma: The experimental results are rarely replicable in the strict scientific sense, but the anomalous yields are well beyond chance expectations and a number of common features thread through the broad range of reported effects. Various attempts at theoretical modeling have so far shown little functional value in explicating experimental results, but have served to stimulate fundamental re-examination of the role of consciousness in the determination of physical reality. Further careful study of this formidable field seems justified, but only within the context of very well conceived and technically impeccable experiments of large data-base capability, with disciplined attention to the pertinent aesthetic factors, and with more constructive involvement of the critical community.

Prologue

"The world of psychic phenomena might be likened to a vast, fog-shrouded swamp, wherein are reported to dwell a bewildering array of bizarre phenomenological creatures, all foreign to our normal perceptual and analytical catalogs. Some scholars who have explored this clouded domain have returned to announce categorically that all such life is illusory—mere sunken stumps and swirling subsurface shadows, inviting misperception by the gullible and misrepresentation by the purveyors. But others of comparable conviction have described in minute detail their observations of a variety of extraordinary beings of awesome dimensions and capability. Some of these are claimed to appear unexpectedly, erupting from the roily depths to flash momentarily in the sunlight of human experience, only to disappear again before any systematic calibration of their characteristics can be taken. Others are reportedly enticed to more replicable and controlled behavior, but only by persons of special talent or extensive training. Much invalid, even fraudulent evidence of such activity has been touted by exploiters of these mysteries, thereby casting deep suspicion on all other testimony. When fully sifted, only a very few legitimate specimens seem to have been captured, by tediously deliberate trolling of the brackish domain, or by more incisive invasion of its turbid interior, and even these have proven so incomprehensible and so delicate to exposure, and the imposed criteria for their credibility have been so severe, that they have not been fully persuasive. Yet the goal remains alluring, and the search continues.

Manuscript received July 15, 1981; revised October 26, 1981. This work was supported in part by the McDonnell Foundation, Inc., and by the John E. Fetzer Foundation, Inc., The Explorers Club, the Institute of Noetic Sciences, and the Little River Foundation.

The author is Dean of the School of Engineering/Applied Science, Princeton University, Princeton, NJ 08544.

Introduction

WITH THIS unlikely bit of allegorical musing, I venture to begin the most extraordinary writing task I have yet attempted: to respond to the request of the Editors of this journal for a critical review of the status and prognosis of scientific research into so-called psychic phenomena. I do so with some trepidation, first because the topic is far from my principal line of scholarship and my involvement with it has been brief and tightly circumscribed, and second, because of the intensity of reactions any commentary on this subject tends to call forth from many quarters.

For these reasons, it may be well at the outset to specify my perspective on the field and the purpose that I hope this article will serve. My formal training is that of an engineer and applied physicist, and the bulk of my research has concerned a sequence of topics in the broad domain of the aerospace sciences: Fluid mechanics, ionized gases, plasmadynamics, and electric propulsion. In my present position as Dean of the School of Engineering and Applied Science of Princeton University, I have occasion to be involved with an even broader selection of topics selected for undergraduate independent projects, and it was in that context some four years ago that I was requested by one of our very best students to supervise a study of psychic phenomena. More specifically, this young lady proposed to bring her talents and background in electrical engineering and computer science to bear on some experiments in controlled, low-level psychokinesis. Although I had no previous experience, professional or personal, with this subject, for a variety of pedagogical reasons I agreed, and together we mapped a tentative scholarly path, involving a literature search, visits to appropriate laboratories and professional meetings, and the design, construction, and operation of simple experiments. My initial oversight role in this project led to a degree of personal involvement with it, and that to a growing intellectual bemusement, to the extent that by the time this student graduated, I was persuaded that this was a legitimate field for a high technologist to study and that I would enjoy continuing to do so.

0018-9219/82/0200-0136$00.75 © 1982 IEEE

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

I have since assembled a small professional staff, secured the requisite funding from a few private sources, and undertaken a modest experimental program in selected aspects of the field that could ultimately have some engineering implications. I should emphasize that my fractional involvement with this program remains quite minor in comparison to my other responsibilities, and that the work is still very preliminary and tentative, but it provides the base of cognizance for my broader observations on the field as a whole.

The intention of this article is to provide some balanced perspective on the modern status of this conceptually and logistically difficult subject. Certainly no field of scholarly endeavor has proven more frustrating, nor has been more abused and misunderstood, than the study of psychic phenomena. Dealing as it does as much with impressionistic and aesthetic evidence as with analytical substance, and carrying by its nature strongly subjective and numenistic overtones, it has been incessantly prostituted by charlatans, lunatics, and sensationalists, categorically rejected by most of the scientific establishment, and widely misunderstood by the public at large. Interspersed with this, and greatly encumbered by it, a pattern of legitimate effort to comprehend and utilize the purported phenomena has evolved to a point where some dispassionate assessment of its accomplishments can be attempted. The questions addressed by this review are whether, once the overburdens of illegitimate activity and irresponsible criticism are removed, there remains sufficient residue of valid evidence to justify continued research and, if so, how this research might most effectively be styled, facilitated, and evaluated.

Before addressing these issues directly, it may be helpful to review briefly the historical evolution of the field, its contemporary nomenclature and conceptual organization, and the dimensions of current activity. This can then be followed by a general overview and critique of the modern research, and that in turn by more detailed description of a few specific efforts, drawn primarily from our own work. Toward the close, we shall attempt to survey several theoretical approaches to modeling of psychic processes and comment briefly on potential implications and applications of the phenomena. In all of this, no tone of advocacy is intended, other than for objective assessment of the evidence in hand.

History

In a sense, the study of psychic phenomena is one of the oldest of human endeavors [1]–[7]. As far back as can be traced, mortal man has pondered the supernatural in one form or another. Cave drawings at Lascaux and Altamira, circa 20 000 B.C., reflect this preoccupation, and the religious rites of early societies of both the eastern and western worlds were heavily loaded with psychic formalisms. The classic civilizations of Egypt, Greece, and Rome dealt extensively in psychic process. The Delphic Oracle was politically important from the earliest Hellenic times to the age of Alexander the Great, and was consulted on problems as diverse as the proper measures to stop a plague, the constitutions of Greek city-states, and the best locations for new colonies. Even Aristotle, one of the most empirical of the classical philosophers, examined the causal links in prophetic dreams.

Virtually every form of organized religion practiced by man has been thoroughly laced with various forms of psychic mechanism. The Bible, like most other basic theological texts, treats psychic process as a central ingredient, in a tone so matter-of-fact that one is inclined to believe that people of those times accepted such events rather routinely. Indeed, the Bible is an excellent catalog of psychic phenomena; virtually every category of effect identified today is illustrated there in one form or another.

Christian writers and philosophers, from Augustine to the Reformation, recount many purported instances of psychic phenomena, usually attributed to visitations of divine grace or demonic posession. Secular medieval writing also abounds with supernatural and mystical reference, and even in the Renaissance period it is still difficult to separate psychic allusion from religious dogma, although both were then translated into more organized forms in art and literature. Early in the 16th century the celebrated Swiss physician and philosopher Paracelsus wrote extensively on psychic capabilities and potentialities. In his words:

> The mind of man is the microcosmic counterpart of the universal mind ....One man may communicate his thoughts to another with whom he is in sympathy, at any distance however great it may be, or he may act upon the spirit of another person in such a manner as to influence his actions .... [8].

Perhaps the first major scientific commentaries on the topic were offered near the turn of the 17th century by Sir Francis Bacon, widely regarded as the originator of the scientific method. In *The Advancement of Learning* he suggested that "superstitions and the like" should not be excluded from scientific study, and in his posthumous book, *Sylva Sylvarum*, he proposed deliberate investigation of telepathic dreams, psychic healing, and the influence of "imagination" on the casting of dice [9]. Some years later, a group of British intellectuals including Henry More and Joseph Glanvill met regularly to discuss paranormal topics, and in 1681 Glanvill published the substance of these studies in a book entitled *Saducismus Triumphatus* [10].

Meanwhile, some four centuries of public and church hysteria over sorcery and witchcraft, as manifested in a sequence of trials, tortures, and executions, had begun to subside, and by the mid-18th century, the Roman Church authorized Prospero Lambertini, who later became Pope Benedict XIV, to carry out a scholarly investigation of reports of psychic events. His conclusions, recorded in *De Canonizatione* [11], were surprisingly unecclesiastical: namely, that 1) psychic experiences were not necessarily divine miracles, but could occur to "fools, idiots, melancholy persons, and brute beasts"; 2) apparitions had little to do with sanctity or demonic entities; 3) prophesy occurs more often in sleep than in waking; 4) it is difficult for a prophet to distinguish his own thoughts from extrasensory messages; and 5) predictions frequently take symbolic forms. In all of these, he presaged to some degree modern thoughts on these topics.

At roughly the same time, Anton Mesmer's discovery of hypnosis opened an alternative route to demonstration and study of unconscious psychic process that has continued to this day. Early reports of hypnotized subjects performing telepathic or clairvoyant tasks were common [12], [13], and although much of this evidence might now be discounted on the basis of inadequate experimental control, interest in hypnosis specifically, and in various altered states of consciousness generally, as facilitators of psychic experience persists into some of the modern experimentation.

Also in this mid-18th century period, a spiritualist movement focused on extrasensory contact with the dead, possibly influenced by the work of Emanuel Swedenborg [14], [15], germinated in this country as well as in England, and by the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

19th century had reached the dimensions of an organized religion. Symbolic of the popular preoccupation with the topic, Mary Todd Lincoln was reported to have held séances in the White House in the early 1860's [16]. A classic two-volume work by F.W.H. Myers, entitled *Human Personality and the Survival of Bodily Death* [17], brought the topic to its acme of sophistication, but eventually the fanaticism the movement attracted and its fraudulent exploitation created a negative attitude in the scholarly community which prevails yet today.

Despite these millennia of human concern with the paranormal, orderly and organized scholarly search for verification and understanding of psychic phenomena began only a century ago, with the establishment in London in 1882 of the Society for Psychical Research, in whose Proceedings appeared the first formal publication of controlled experiments in telepathy and clairvoyance [13], [18], [19]. Three years later the counterpart organization in this country, the American Society for Psychical Research, was founded in Boston by several distinguished scientists and philosophers. Because of financial difficulty, this shortly merged with the British group, but reemerged in 1905 as a separate entity with its own professional journal, and has continued as such to the present [20].

Although the SPR attracted a barrage of criticism from the scientific and intellectual communities, it also attracted significant participation of eminent scholars from established fields. Numbered among its presidents are three Nobel Laureates, ten Fellows of the Royal Society, one Prime Minister, and a substantial list of physicists and philosophers, including Henry Sidgwick, Frederic W. H. Myers, Lord Rayleigh, Sir J. J. Thomson, William McDougall, Edmund Gurney, Sir William Crookes, Sir William Barrett, Henri Bergson, Arthur, Earl of Balfour, Gardner Murphy, G.N.M. Tyrell, Charles Richet, Gilbert Murray, and one of the most articulate contributors to the evolution of critical thought on this topic in this period, the Harvard psychologist and philosopher, William James. One of the founders of the ASPR, James wrote extensively and eloquently on behalf of objective and disciplined study of psychic phenomena [21]-[25]:

> Any one with a healthy sense for evidence, a sense not methodically blunted by the sectarianism of 'science,' ought now, it seems to me, to feel that exalted sensibilities and memories, veridical phantasms, haunted houses, trances with supernormal faculty, and even experimental thought-transference, are natural kinds of (phenomena) which ought, just like other natural events, to be followed up with scientific curiosity [25].

Entering the 20th century, a new perspective on psychic phenomena was provided by the emergence of psychology as a scholarly discipline, and especially by the early efforts in clinical psychology and psychoanalytic therapy. The patriarch of this evolution, Sigmund Freud, was a member of the SPR and contributed, albeit somewhat reluctantly, to its publications [26], [27]. His recognition and exploration of the unconscious mind and of the function of dreams prompted Myers to suggest a possible explication of various psychic effects which is still of theoretical value [17]. Freud's interest in parapsychology increased toward the end of his life, and he is reported to have conceded informally that were he to begin his career anew, he would focus on this topic.

Freud's former protegé, Carl Jung, who had written his Ph.D. thesis on the psychology of occult phenomena, pursued exploration of the unconscious to deeper dimensions of paranormal experience, publishing widely on such subjects as telepathy, mediumship, synchronicity, the collective unconscious, and theoretical models of psychic process [28]-[30]. In *Memories, Dreams, Reflections*, he stated:

> ... the relationship between doctor and patient, especially when a transference on the part of the patient occurs, or a more or less unconscious identification of doctor and patient, can lead to parapsychological phenomena. I have frequently run into this [30].

Jung's collaboration with the eminent physicist Wolfgang Pauli on the topic of synchronicity clearly influenced the subsequent evolution of both careers and of fundamental concepts in both disciplines [31]. Although much of the established psychological community has since rejected parapsychology as a valid discipline, some interest has been retained by a few clinical practitioners, presumably because of the demonstrated concomitance and similarities of apparent psychic experiences with certain psychological processes [32], [33].

It was also early in this century that the first organized academic studies of psychic phenomena were mounted. One of the more visible of these devolved from gifts and a bequest from Thomas W. Stanford, brother of the founder of Stanford University, to endow psychic research at that institution, and to this day the university provides support of a "psychic research fellow" and retains a collection of so-called "apports" indicative of the donor's long personal involvement with the field. Modest research programs were also undertaken at Harvard and a few European universities in the first decades of this century, as evidenced by occasional publications in various established journals.

The benchmark academic effort, however, germinated at Duke University in the late 1920's, when William McDougall, who had been James' successor at Harvard, arrived to chair the department of psychology and appointed J. B. Rhine and Louisa Rhine, "to study the claims to scientific value of the field known as psychical research." Their early tentative efforts in the study of postmortem survival gradually evolved into a laboratory for controlled research in "extrasensory perception," as they first termed the process. In this laboratory were established many of the basic concepts and protocols of modern psychic research, as well as the first extensive and systematic data bases of several types of psychic experimentation. The professional and personal history of the Rhines and their laboratory is a fascinating saga in its own right, but would take us too far afield here [34]-[37]. A few excerpts from a 1967 address of J. B. Rhine to the American Psychological Association, in which he attempted to summarize his first two decades of intensive study, give hint of the inherent attractions and frustrations of this field, and of the man's optimistic vision:

> The phenomena that were being studied began to show lawful interrelations and even a degree of unity. One by one the major claims, based originally only upon spontaneous human experiences, were subjected to laboratory test and experimentally verified.... Certain general characteristics of the psi process became clear during this period. The most revealing of these is the subject's lack of conscious control over any type of psi ability, a characteristic which accounts for its elusive nature. It was new methodological ground, even for psychology.... Also, we were surprised to find that psi ability is widespread, probably even a specific human capacity rather than a capability possessed by a few rare individuals as had been the popular belief. Evidence that psi is not linked with illness or abnormality was another welcome advance.... By 1951... a healthy young science was emerging [38].

In 1937, the Rhines began publication of the *Journal of Parapsychology*, which remains a leading journal in the field today. A professional organization calling itself the Parapsychological Association was formed in 1957, and in 1969 was

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

accepted as an affiliate by the American Association for the Advancement of Science.

At the present time, there are eight English language publications covering this field [39], supplemented by numerous less formal magazines and countless books of widely varying quality and relevance. Research activity is reported from some twenty U.S. universities and colleges and at least as many institutions in Western Europe [40], but in most cases it is of very small scale. There are very few academic programs of study, although some fifty M.A. and Ph.D. theses have been accepted on psychic topics at reputable universities over the past forty years [41]. Some ten research institutes and private corporations in the United States have also authorized publications and reports in the field [42]. The extent of Eastern Bloc and Oriental efforts [43]-[54] and of classified research in this country are matters of considerable speculation on which I cannot comment with authority.

Further review of contemporary programs will be attempted in subsequent sections, following an outline of modern nomenclature and conceptual organization of the topic. In closing this historical overview, we might simply observe that in many respects the growth pattern of this field resembles that of the natural sciences in their earliest days, or perhaps even more the incubation of classical psychology, in terms of the absence of replicable basic experiments and useful theoretical models, the low level of financial support and internal professional coordination, and the low credibility in the academic establishment and public sectors. Also like those fields, the survival and early growth of psychic research can largely be attributed to the efforts of a few scholars of sufficient conviction, stature, and courage to withstand the rejection of the orthodox communities.

Nomenclature and Conceptual Organization

Before turning to an assessment of contemporary research, it may be useful to specify some notation and delineation of the field, to an extent consistent with the present limited comprehension of the phenomena. First, let us agree to a global definition of "psychic phenomena" (frequently denoted by "psi" or "ψ") to include all processes of information and/or energy exchange which involve animate consciousness in a manner not currently explicable in terms of known science. Similarly, let "psychic research" imply any scholarly study of such phenomena employing scientific methodology, as opposed to any dogmatic, ritualistic, or theological approaches. Within these definitions, the field may then be roughly divided into two major categories: extrasensory perception (ESP) and psychokinesis (PK).

ESP refers generally to the acquisition of information from sources blocked from ordinary perception. Under this category are included such subdivisions as telepathy, which refers to detection of another person's thoughts; clairvoyance, which refers to contemporary perception of remote physical objects or events; precognition and retrocognition, which refer to perception of future events and events in the past not accessible by normal recollection; and animal ESP, which encompasses a variety of seemingly inexplicable capabilities, such as homing, psi-trailing, collective behavior, communication, etc.

PK (occasionally termed telekinesis, or psychoenergetics) refers to a palpable influence of consciousness on a physical or biological system. The interaction may be deliberate or spontaneous, and the energy transfer involved may range from microscopic disturbance of atomic-level processes, through macroscopic distortion or levitation of objects, up to some very drastic "poltergeist" effects. Psychic healing and man-plant interactions would be two examples of PK in biological systems.

Note that in its major subdivision into ESP and PK, the field conforms to two of the main categories of present-day science and high technology, i.e., that encompassing the extraction, conversion, transmission, storage, and utilization of information, and that comprising the same sequence of processing of energy. Other domains of psychic research can be identified which do not fit into these major categories of ESP and PK and with which we shall not be further concerned in this article. Examples would include research into survival of death, and the family of "out-of-body experiences (OBE)," including astral projection, autoscopy, and bilocation. The following table attempts to summarize the subdivisions in a concise form.

Categories of Psychic Phenomena

- I. Extrasensory Perception (ESP)
  - A. Telepathy
  - B. Clairvoyance
  - C. Precognition/Retrocognition
  - D. Animal ESP
- II. Psychokinesis (PK)
  - A. Physical Systems
  - B. Biological Systems
- III. Survival
  - A. Reincarnation
  - B. Apparitions
  - C. Mediumship
- IV. Out-of-Body Experiences (OBE).

Clearly this particular arrangement is neither unique nor orthogonal. Elements of one category frequently appear in the context of another, e.g., precognitive clairvoyance; telepathic healing, etc., and occasionally an assignment is ambiguous, e.g., a particular effect may be regarded as precipitated by PK, or simply to be forecast by precognition. Notwithstanding, the table may aid in keeping the subsequent illustrations in some order.

Pattern of Contemporary Research

By its nature and heritage, modern psychic research remains rather diffuse and lightly structured, making any attempt to catalog the work by institution or laboratory, or by tracing developments of given lines of effort, rather ineffective and premature. Instead, it may be more useful to comment on the pattern of attention to this field by academic disciplines, noting the variations in emphasis, methodology, representation, and interpretation brought to bear, using specific projects only as illustrations with no implications of hierarchy or attempt at completeness. Even in this format, no recitation of specific research results or conclusions will be attempted, since these can be misleading or incomprehensible when extracted from the detailed context of their experimental arrangements and protocols. In later sections, an effort will be made to follow a few sample experiments through to their particular results and conclusions.

By far the most sustained and broadcast attention to this field has been given by a cadre of scholars with professional backgrounds in classical psychology, comprising a controversial subdiscipline termed "parapsychology." This group has tended to approach the field with the traditional psychological protocols and vocabulary, and to interpret results in the context of

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

their clinical, cognitive and behavioral psychological heritages, with the natural consequence that they have concentrated mainly on the ESP category of psychic tasks, although some PK work dots their recent literature. Perhaps the most extensive class of parapsychological research has attempted to correlate psychic performance with personality variables. The age, sex, creativity, openness, hostility, extroversion, motivation, and intelligence of the participants as indices of ability to perform ESP tasks have been explored at length, and some significant correlations, most notably with positive *a priori* attitudes toward the tasks and with outgoing, creative personalities, have been reported from several laboratories. Other studies have searched for connections between psychic performance and dream recall, learning and response strategies, memory, and feedback [55]-[61].

A more aggressive style of parapsychological research has invoked a variety of altered states of consciousness in attempts to enhance psychic process. These have included various natural and traditional practices, such as sleep, meditation, and progressive relaxation [62]-[67]; more mechanical sensory inhibition strategies such as hypnosis, isolation and "ganzfeld" [66], [68]-[70]; and a few controversial efforts with drug-induced states [71]. Physiological correlates have also been sought, using conventional EEG, GSR, and plethysmographic equipment to monitor neurological, cardiovascular, and muscular response to psychic effort [35], [72]-[75]. The difficulty of obtaining successful replications of previously positive results and an observed common tendency of participant performance to deteriorate over time ("decline effect") have led to systematic study of the role of the experimenter in eliciting results, i.e., to consideration of the possible influence of the experimenter's personal attitudes, expectations, and style of interaction with his subjects, as well as the overall environmental ambience of his laboratory, on the experimental yield [76]-[81].

Despite its present recalcitrance toward more systematic study of psychic phenomena, the richly diverse, rapidly maturing parent field of psychology continues to offer an expanding array of modern methodologies and models which could be brought to bear on increasingly sophisticated study of this subject. Computer-assisted linguistic analyses; psychoneurological studies of attention, perception, and concept formation; social learning theory and similar approaches to human interactions; and the emerging formulations of transpersonal and humanistic aspects of human consciousness, all bear possible relevance to comprehension of various aspects of this ultra-difficult stepchild, but at the moment, the low level of financial support, and negative professional peer pressures have discouraged such enterprise.

The involvement of physicists in psychic research, while considerably less extensive, has been no less dedicated and no less controversial. Since the days of Sir Francis Bacon, a number of noted physicists have made excursions into this field, usually to the bemusement and ridicule of their colleagues of the day. One of the most notable of these was Sir William Crookes, discoverer of the element thallium and pioneer in the physics of low pressure discharges, whose broadside professional and personal battles with the scientific establishment over this issue make entertaining, and possibly enlightening, reading [82]. Sir Isaac Newton was intensely involved in the study of alchemy, including some of its more metaphysical aspects [83], and as already mentioned, Lord Rayleigh and J. J. Thomson were active members of the S.P.R.

In more recent years the attention of physicists has influenced development of the field in three ways. First, their interests have focused more on the PK category of problems, i.e., the interaction of human consciousness with physical systems, to balance the predominant ESP interests of the parapsychologists. Second, more sophisticated experimental equipment than has typically been available to the psychological community has been brought to bear on the identification and correlation of very low-level physical effects. Third, the traditional theoretical physics formalisms have been directed to the proposition of various models of psychic phenomena, from whence has arisen some hope of establishing the traditional dialogue between critical experimentation and theoretical hypothesis essential to any ultimate comprehension and application of such phenomena.

Typical of the modern physicist's specific contributions to the field have been the development and application of a variety of electronic random event generators (REG) for the purpose of identifying and correlating PK abilities in human subjects [84]-[93], and similar application of magnetometers [94], torsional pendula [95], lasers [96], interferometers [97], and electronic strain gauges [98], [99] to a variety of other PK tasks. On the theoretical side, a number of applications of quantum mechanics, statistical thermodynamics, electromagnetic theory, and other formalisms to the representation of psychic process have been proposed [100]-[113], and attempts at some philosophical correlation of the phenomena with other previously or presently obscure physical processes have been suggested [114]. Again, despite the open identification of a few distinguished personalities with such efforts, a more broadly held categorical rejection of the field has inhibited much collaborative or systematic attention to it.

Up to this time, the involvement of engineers with psychic research has been very recent, very sparse, and very much along the lines of the experimental physicists. Beyond our own program, which will be outlined in some detail below, I am aware of only a very few engineering laboratories addressing any aspects of the field in any substantial and deliberate way [115]. These have so far tended to concentrate on applied physics types of experimentation and on aspects of information processing, rather than on more empirical technological applications.

Another community of scholars to influence the pattern of psychic research comprises the statisticians and other applied mathematicians and logicians who have been concerned with the proper evaluation and interpretation of the research data. In the absence of any experiments displaying rigid causal replicability, all of the inferences and hypotheses about psychic phenomena have necessarily been based on either anecdotal or statistical evidence. The former defy any systematic representation; the latter are vulnerable to alternative interpretations and hence to impressionistic bias and argument.

Early in the emergence of mathematical statistics as an integral discipline, S. S. Wilks found himself involved in a controversy over the validity of the statistical procedures of early psychic researchers, and published some recommendations for methods that could be applied to telepathy experiments [116]. Since that time, much of the commentary from the critical community has addressed perceived flaws in the statistical methodology underlying the experimental evidence [117], [118], and the advocate community has reacted by paying disproportionate attention to this aspect of their logic. Most of the encyclopaedic references in the field contain

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

substantial components on the statistical methods [119], and the leading journals regularly display intense dialogues on specific statistical issues [120]-[124]. At least one of these journals routinely refers all articles submitted for publication to a consultant statistician as a part of their review process. A few illustrations of the statistical questions that can arise in psychic experiments appear in the detailed examples presented below.

A number of other disciplines have played roles in the evolution of the study of psychic process and continue to contribute, albeit somewhat more peripherally. A succession of philosophers, from Aristotle through James and Bergson to C. D. Broad in the present era, have mused on these topics [21]-[25], [125]-[129]. The intersections of the field with anthropology, theology, and history have been approached from many perspectives ranging from aesthetic to analytic, and from dogmatic to scholarly [130]-[134]. Its relevance to the study and practice of medicine has been an enduring and intense debate, focusing in the present day on the propriety and efficacy of holistic health strategies and psychic healing. Isolated instances of interest on the part of chemists, biologists, geologists, and archaeologists can be found, and the application of psychic techniques in criminology and law enforcement, while less rigorous than most of the academic efforts, contributes further anecdotal evidence to the overall data base. In the arts and humanities, the topic continues to provide stimuli for a variety of creative compositions.

Still other areas could be cited, but these become progressively more satellite than central to the task of this paper. Rather than pursuing this disciplinary survey further, it seems preferable next to focus in greater detail on a few contemporary studies that can serve to illustrate more specifically the bizarre phenomena, the awkward and tedious protocols, and the unconventional theoretical concepts that arise in this class of research. Although these will provide a better sense of the status of the field than any attempt to summarize results from the diffuse multidisciplinary pattern of effort sketched above, one general assessment may be useful at this point. Namely, throughout all of the work just skimmed, and that sharpened somewhat below, I am aware of no reputable investigator who has claimed, let alone demonstrated, any psychic experiment approaching classical scientific replicability. What have been put forward are a varied assortment of observations, currently inexplicable in terms of established science, which display certain common phenomenological and psychological features, and which could have substantial implications for basic physical theory and ultimate practical applications. The following examples are presented in that spirit.

Psychokinesis

The first group of experiments selected as more detailed illustrations of contemporary psychic research are drawn from the general subdomain of PK. As defined above, this broad category of purported psychic phenomena encompasses the possible influence of human consciousness on the behavior of physical or biological systems or processes, and comprises several loosely related classes of effect characterized by different scales of energy, forms of manifestation, replicability, and statistical behavior. Confining attention to interactions with physical objects or systems, the most popularly publicized class features the deformation, levitation, or other macroscopic disturbance of objects, as commonly propounded by professional performers, mediums, and various Eastern practitioners [135]-[137]. Although a number of serious efforts have been made to submit such demonstrations to rigorous scientific testing, these have tended to yield only equivocal confirmations, fodder for the critical community, and some embarrassment and frustration to the investigators.

Of a yet more bizarre nature are the family of very rare and spectacular "poltergeist" effects, more technically termed "recurrent spontaneous psychokinesis" (RSPK), wherein are reported specific major events of levitation, vibration, teleportation, and breakage of a wide range of objects, a variety of acoustical and electromagnetic phenomena, and various optical aberrations [138], [139]. For years these phenomena were naively attributed to manifestations of the spirit world, or return of the dead to "haunted" houses, and inspired countless horror movies and pulp-magazine articles. Recently, some order has been brought to this weird business by systematic surveys of documented poltergeist cases undertaken by A.R.G. Owen, W. G. Roll, J. G. Pratt, and others [138]-[142]. In one of these surveys, 116 cases of reported poltergeist activity, ranging back to the year 1612, were re-examined. Of these, 92 were found to be associated with particular individuals living in the affected dwelling, most of whom were adolescents, and most of whom were affected by some neurological/emotional ailment, most commonly epilepsy. Often a precipitating traumatic event could be identified which seemed to initiate the activity, after which the general pattern involved a period of relatively mild precursor events, a sequence of major disturbances, and a period of "after shocks," extending as much as several weeks beyond the main events. Controlled experimentation on poltergeist phenomena has proven virtually impossible because of their infrequent and unpredictable occurrence pattern and because of the delicate physiological and emotional situations prevailing in most cases, which have taken precedence over the technical investigations. Nevertheless, these processes have retained some fascination because of the magnitude of the energy transfer involved and because of their apparent correlation with specific types of individuals and neurological disorders.

The most systematic and persuasive studies of PK, however, have dealt with much more modest scales of physical disturbance, in some cases reaching down to the atomic level. This somewhat more viable domain has been addressed by numerous investigators in various ways, but basically one of two strategies is followed. In one approach, relatively simple physical systems are employed—mechanical, electrical, optical, thermal, etc.,—each of which involves a particular component or process that is ultra-sensitive to disturbance. The experiment is arranged to signify such disturbance by a relatively large change in some display which provides feedback indication to the operator, much in the spirit of a biofeedback instrument, and simultaneously to provide some form of permanent data record. Examples of this class of experiment would include the use of magnetometers, torsional pendula, optical interferometers, electronic strain gauges, glow discharges [143], and sensitive thermistors [144].

In the second approach, attempt is made to distort the normal statistical patterns of various random physical processes on either a microscopic or macroscopic scale. In a sense, these experiments deal with energy rearrangement within the systems, i.e., with their information content or entropy, rather than with energy transfer to the system *per se*. The earliest versions of this class employed dice, or other simple mechanical implements of well-known statistical behavior [37], [92], [145]-[148], but more recent studies have tended to employ more sophisti-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

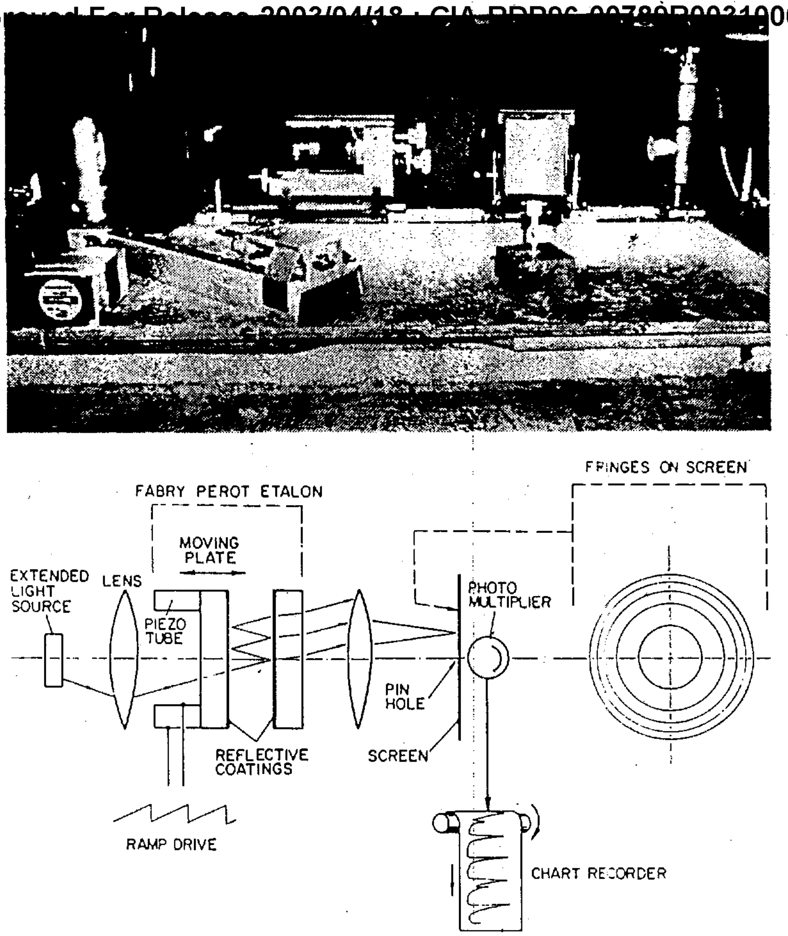


Fig. 1. Photograph and schematic of Fabry–Perot interferometer.

cated apparatus, such as electronic REG's like those described in detail below.

Although the first two categories of PK feature more dramatic effects and thus have enjoyed greater popular attention, the smaller scale phenomena seem more amenable to controlled experimentation and theoretical interpretation, and the remainder of this section deals only with such. Perhaps the most efficient means of elaborating on this type of research would be to review briefly the spectrum of such experiments in progress in our own laboratory, and then to display and attempt to interpret data from one of them. In so doing, we intend no neglect of other work noted in the references, but simply deal from greater familiarity.

As examples of the first class of low-level PK experiment mentioned above, we have in operation a Fabry–Perot optical interferometer, a dual-thermistor bridge, and a photoelastic strain detector. In the interferometer experiment, shown in the photograph and schematic of Fig. 1, a Coherent Optics Instrument #360–370, using a diffuse sodium lamp as source, is adjusted to produce circular fringes on a screen visible to the operator (Fig. 2). Small changes in the separation of the interferometer plates cause the fringes to migrate radially inward or outward. By visual observation of the fringe movement, plate motions of less than 0.1 wavelength can be readily detected. Via a pinhole in the screen, the brightness of the central fringe is recorded by a photomultiplier/chart recorder system at an order of magnitude higher sensitivity, thereby preserving quantitative output data while the operator simultaneously sees an attractive optical display of his progress for use as feedback.

The task of the operator is to elicit significant migration of the fringe pattern in a stated direction relative to the normal baseline drift of the instrument. The protocols involve rigid control and monitoring of the environment of the instrument and surrounding laboratory, and the interspersing of baseline responses with active PK efforts obtained under otherwise identical conditions, including the position of the operator and any other personnel relative to the instrument. In pilot studies with this device, a variety of fractional-fringe responses were observed, using several different operators and various initial interferometer settings. A more formal procedure has since been developed which provides more precise conditions for an ongoing series of trials. In this protocol, the central fringe is set initially on a maximum gradient position between a bright and dark fringe, and its progress monitored for subsequent periods of baseline or PK effort. Encouragingly replicable data have been obtained from a number of different operators, in the form of chart recordings of 5-min PK trials with interspersed 5-min baseline drifts of the instrument. Using computerized graphic, regression, and spectral analyses of the data, it is possible to discern characteristics in the hierarchy of trace derivatives and the Fourier spectra which, while not definitive, display certain recurrent features [97]. No physical interpretation has been attempted other than to acknowledge that the observed fringe migrations could also be indicative of slight changes in the index of refraction of the air in the plate gap or in the wavelength of the light source, as well as of a displacement of the plates.

The dual-thermistor experiment comprises a much more sensitive version of a multiple-thermistor arrangement on which PK influence was originally reported by Schmeidler [144].

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

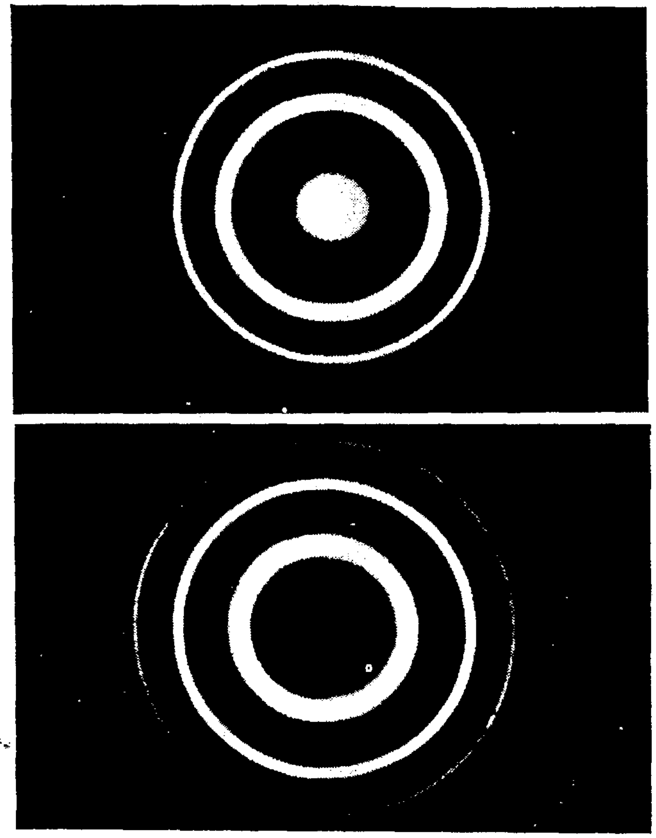


Fig. 2. Interferometer fringe patterns.

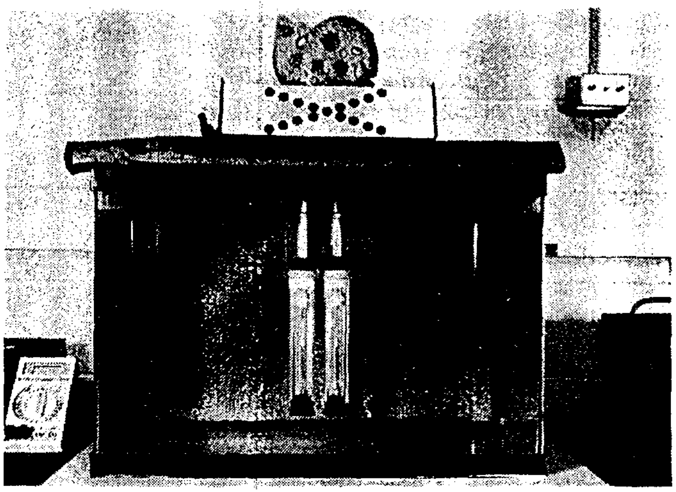


Fig. 3. Dual-thermistor apparatus.

As shown in Fig. 3, two Omega Engineering thermistors, Model UVA 3254, each with its own electronic bridge and voltage source, are connected differentially to a Tektronix 1A7A oscilloscope preamplifier and to a visual feedback display. With suitable ground planes and cable shielding, sensitivities greater than 0.001 K are obtainable, and by subtracting the two output signals the major portion of spurious electrical and mechanical interference is eliminated. The effects of ambient thermal variations in the laboratory are essentially excluded by enclosing each of the thermistors in identical Pyrex flasks immersed in a large liquid reservoir, in which configuration the undisturbed system exhibits stable baselines over long periods of time. Using the same interspersed baseline protocol as in the interferometer experiment, the task of the operator is to achieve an increase in the reading of one thermistor with respect to the other or in some more subtle fashion to alter the PK response relative to the baseline. Some such effects have indeed been observed, but little systematic data have so far been accumulated on this experiment.

Also in a preliminary stage is an experiment to monitor internal strain in a solid specimen via photoelastic optical techniques. Several studies have been reported on the PK deformation of solids, but most of these have employed conventional engineering strain gauges or microacoustic sensors as detectors [98], [99], [135], [149], [150], both of which require substantial interface electronics before a feedback signal reaches the operator, leaving unclear the role of the sensor in any possible PK influence. Although less-sensitive than the electronic methods, photoelastic techniques have the advantage of relating the operator more directly to the sensitive element of the experiment via an attractive optical fringe pattern much like that of the interferometer (Fig. 4). This equipment and technique may also be applied to a sensitive levitation experiment wherein the object is suspended on a photoelastic lever arm of suitable dimensions.

Within the second category of low-level PK experiment, we are employing or are now constructing several devices based on random physical processes, some macroscopic in scale, others deriving from atomic-scale processes. The largest of these involves a 6 X 10 ft apparatus, shown in Fig. 5, which drops some 10 000 $\frac{3}{4}$ in spheres through a "quincunx" array of 336 nylon pegs in about 12 min. As a consequence of the multitudinous collisions with the pegs and with each other, the spheres are dispersed into a good approximation of a Gaussian distribution as they fall into 19 collecting compartments at the bottom. The goal of the operator is to distort the distribution in some prescribed fashion to a significant degree compared to empirical baseline experience. Photodiode counters mounted in funnels at the entrance to each bin provide real-time digital displays of the bin populations to supplement the more qualitative feedback of the growing ball stacks seen by the operator and to provide quantitative data for on-line statistical analysis. Fig. 5 shows a typical baseline distribution for this device and a distorted distribution obtained in a particular PK effort. Full statistical analysis of the significance of any particular achieved pattern is a challenging problem in its own right, since it must deal with a combination of 19 bin populations, each of which has its own empirical baseline mean and standard deviation, all constrained by total ball count.

A somewhat similar experiment, not yet refined, employs a device which allows small metallic or dielectric spheres to bounce on an optically flat, precisely horizontal circular plate of glass, which is oscillated by a vibration coil at frequencies from 10 Hz to 20 000 Hz. In the absence of any external disturbance, a sphere started at the center of the plate executes a random walk toward the outside edge, arriving with equal probability at any azimuth. Since the sphere may make as many as 10^5 collisions in the process, it is vulnerable to statistical distortion of its trajectory and consequent terminus. The task of the operator is preferentially to direct the sphere to a prescribed terminal quadrant.

In an attempt to intervene with a random physical process at the atomic level, we have constructed a large glow-discharge device whose luminous patterns are indicative of the mean free path of the current-carrying electrons against inelastic excitation collisions with the background gas. This device, shown in Fig. 6, presents a 36-in X 2-in diam cylindrical glow marked by a sequence of bright and dark zones along its positive column

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

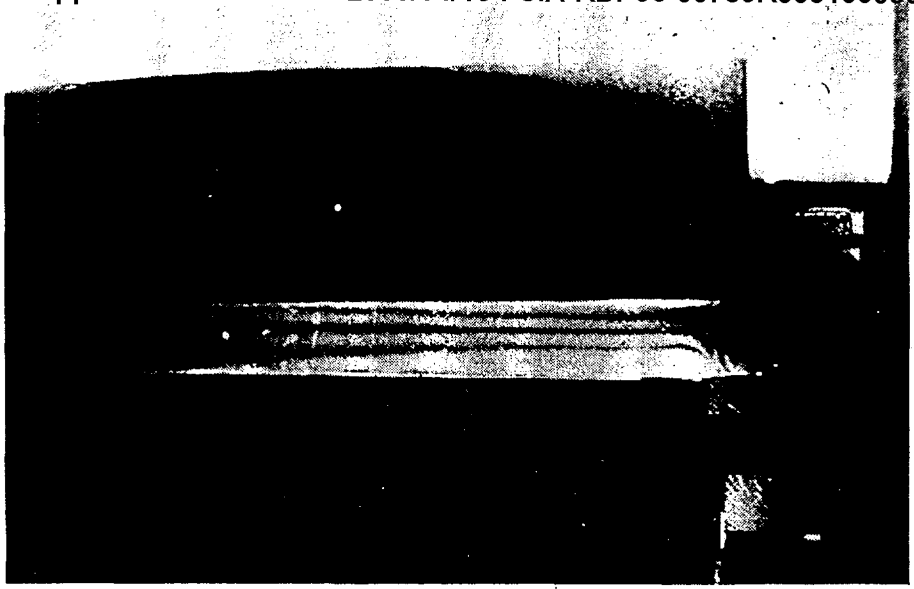


Fig. 4. Photoelastic stress pattern.

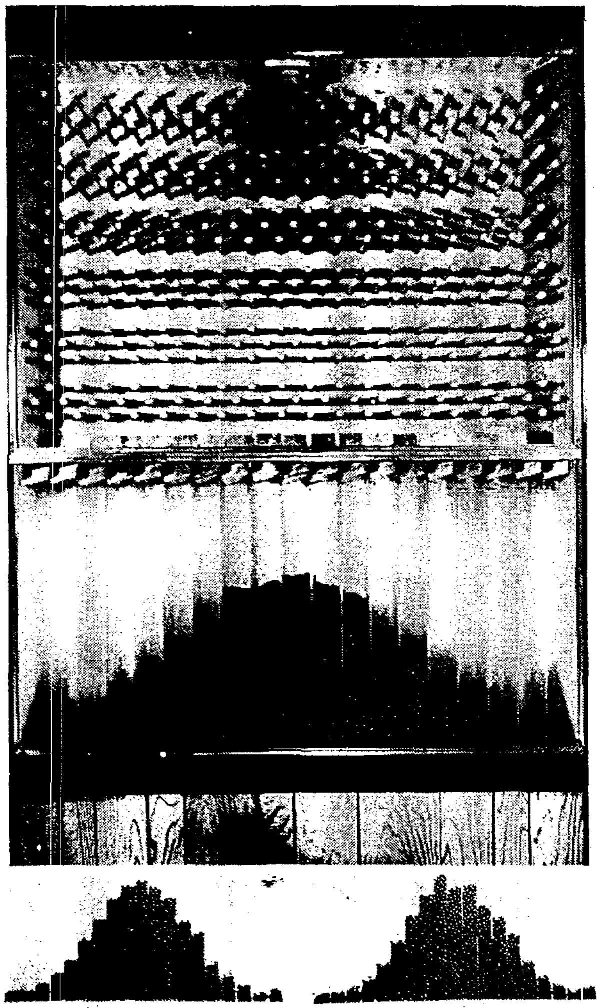


Fig. 5. Gaussian analog device and distributions.

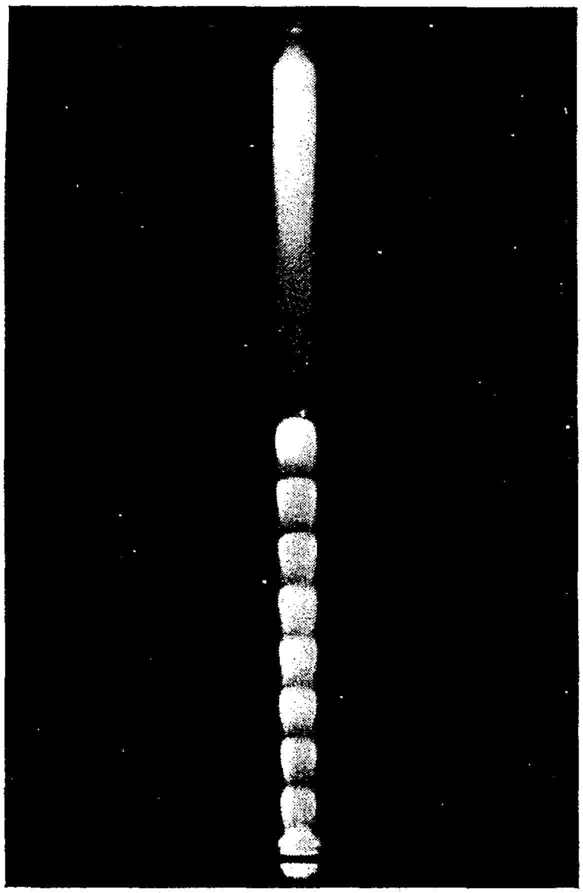


Fig. 6. Glow discharge experiment.

typical of dc discharges in a given range of gas pressure and terminal voltage. The number and locations of these striations are sensitive to the electron inelastic mean free path, which in turn depends on the gas type and density, the electron temperature, and the local electric field. Striation position is monitored by photoelectric detectors, and the goal of the operator is to expand or contract the pattern on demand, to a significant extent compared to the normal background jitter and drift. Protocols are much the same, output data take the same general form, and are analyzed by the same algorithms as in the interferometer and photoelastic experiments.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

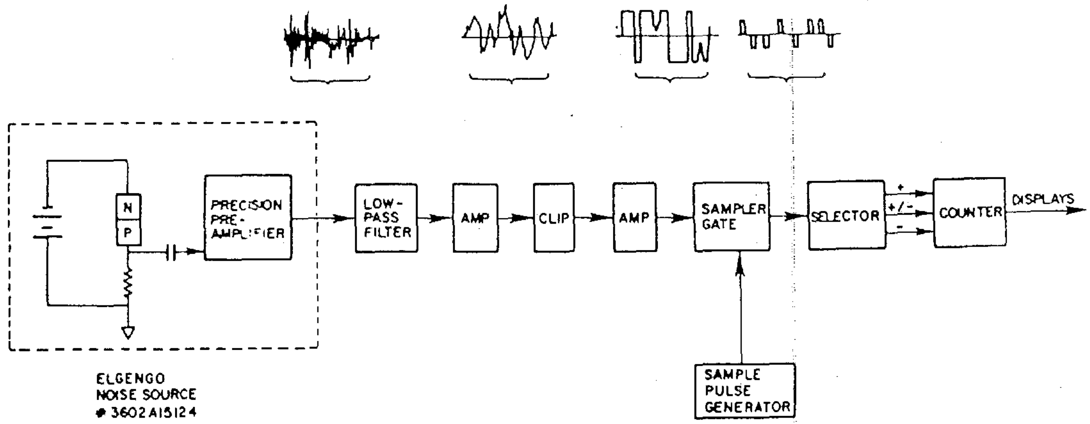


Fig. 7. Functional diagram of REG.

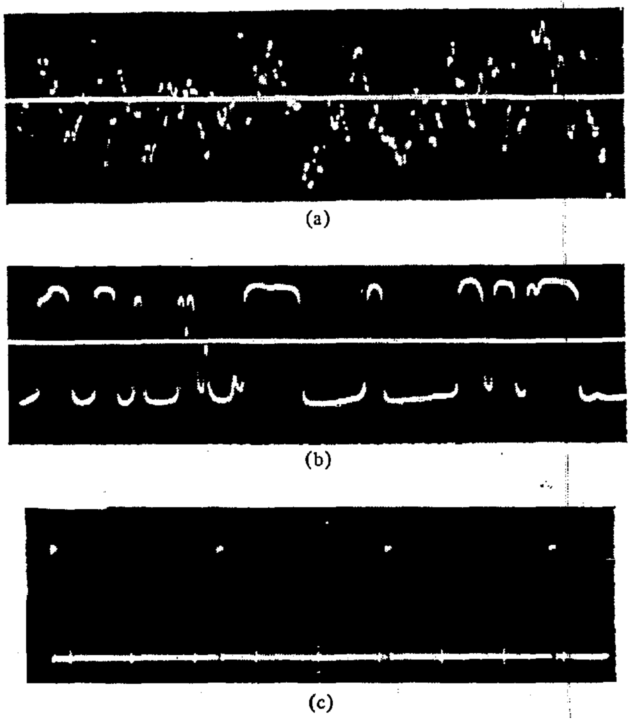


Fig. 8. REG waveforms: (a) Filtered noise. (b) Clipped noise. (c) Sampling pulses.

A number of other atomic-scale random system PK experiments are under consideration, design, or construction, involving such processes as information storage on a microelectronic chip, the spontaneous decay of phosphorescent surfaces, laminar to turbulent transition in a fluid stream, atomic and molecular resonators, and resonant acoustical or electrical cavities, but none of these is far enough advanced to merit description here. Rather, we shall concentrate for the remainder of this section on a more detailed presentation of our most serviceable experiment, and the one on which we have the largest data base, the electronic REG.

Random Event Generator Experiments

REG's have been the most widely used and most productive facilities for experimentation with low-level PK. Although a broad variety of such devices exist, most involve four conceptually and functionally separable components: an electronic noise source; a sampling system which examines the noise at prescribed intervals and prepares an output pulse train corresponding to the samples thus obtained; a system which analyzes the pulse train in accordance with preset instructions and prepares suitable output for a feedback system; and the feedback display itself, which informs the operator of the results of the sampling process.

The particular version we have employed utilizes a packaged commercial noise source module based on a solid-state junction and precision preamplifier (Elgenco Model 3602A15124), but modules employing radioactive decay units or glow discharges can be readily substituted. This source produces a random noise spectrum up to several megahertz, which our logic circuit first filters to a flat spectrum from 50 to 20 000 Hz, then amplifies and clips to the flat-topped profile shown in Figs. 7 and 8. This is then sampled by a regular train of gate pulses, yielding a corresponding random succession of positive and negative output pulses indicative of the sign of the noise at the time of sampling, and these are then counted. Since the average time between

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

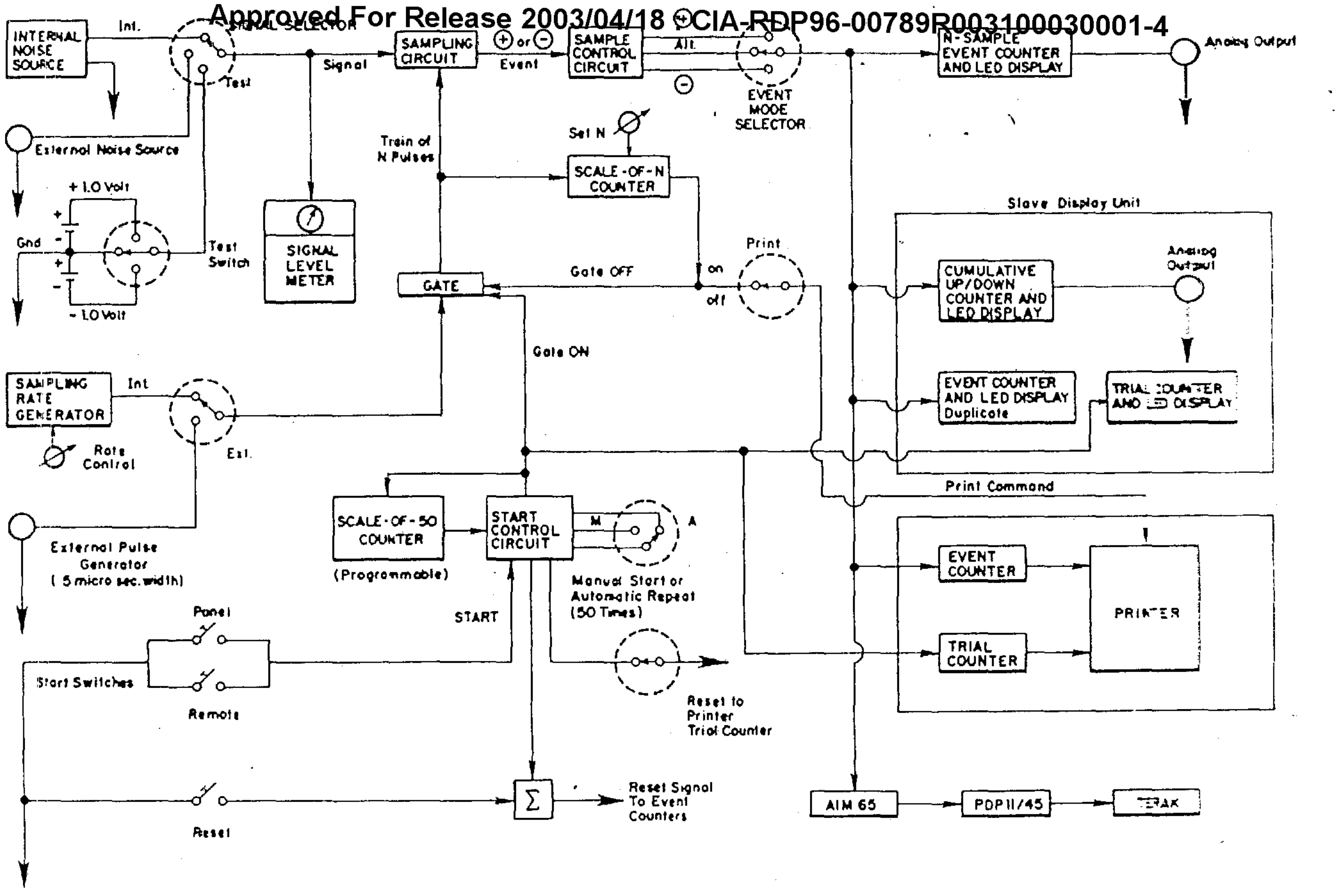


Fig. 9. Electrical schematic of REG.

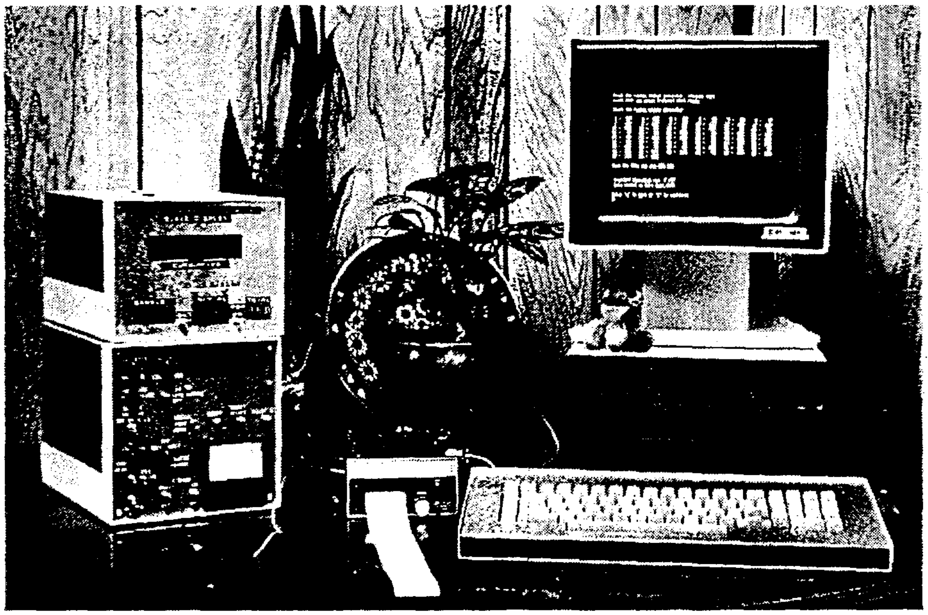


Fig. 10. REG arrangement.

zero crossings of the clipped noise waveform is about 30 μs, sampling rates to about 15 000/s can be tolerated with statistical independence.

The full functional array is sketched in Fig. 9, and a photograph of the boxed units in Fig. 10. By panel setting the sampler may be instructed to take "trials" of 100, 200, or 2000 samples, at a frequency of 1, 10, 100, 1000, or 10 000/s. The counting system may be set to count only positive pulses, only negative pulses, or to alternate positive and negative counting on successive samples. The alternating positive/negative mode effectively factors out any systematic bias in the noise source, and is the mode employed in all the experiments reported here. The counting results are displayed by LED arrays tracking both the running count of each trial and the concurrent mean rela-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE I
REG 200-SAMPLE DATA SUMMARY

Series	Instr.	No. Trials	Mean	Std. Dev.	t-score	P_t	n_+/n_-
REG I	B.L.	12 000	100.009	6.994	0.144	0.443	5678/5611
	PK⁺	4 550	100.264	7.037	2.528	0.006	2230/2056
	PK⁻	3 850	99.509	7.063	−4.313	10^{-5}	1716/1926
	Δ PK	8 400			4.890	5×10^{-7}	
REG II	B.L.	2 500	100.033	6.875	0.239	0.406	1188/1179
	PK⁺	1 950	100.247	6.849	1.590	0.056	916/919
	PK⁻	1 800	99.597	6.775	−2.526	0.006	797/902
	Δ PK	3 750			2.920	0.002	
REG III	B.L.	3 500	99.977	7.013	−0.193	0.424	1658/1655
	PK⁺	2 400	100.227	6.821	1.634	0.051	1150/1086
	PK⁻	2 600	99.736	7.026	−1.918	0.028	1192/1270
	Δ PK	5 000			2.507	0.006	
Σ REG I	B.L.	18 000	100.006	6.981	0.115	0.454	8524/8445
II	PK⁺	8 900	100.250	6.938	3.403	3×10^{-4}	4296/4061
III	PK⁻	8 250	99.600	6.989	−5.203	10^{-7}	3705/4098
	Δ PK	17 150			6.107	5×10^{-10}	
REG Ia	no B.L.						
	PK⁺	2 150	100.206	7.091	1.340	0.088	1059/993
	PK⁻	2 100	99.945	6.937	−0.365	0.358	954/1019
	Δ PK	4 250			1.213	0.113	
REG IIa	B.L.	5 000	100.186	6.974	1.882	0.030	2367/2337
	PK⁺	2 000	100.117	7.041	0.746	0.228	955/950
	PK⁻	1 750	99.941	6.898	−0.360	0.359	803/839
	Δ PK	3 750			0.772	0.220	
Σ REG I	B.L.	23 000	100.045	6.980	0.978	0.164	10891/10782
Ia							
II	PK⁺	13 050	100.223	6.979	3.644	10^{-4}	6310/6004
IIa							
III	PK⁻	12 100	99.709	6.968	−4.596	2×10^{-6}	5462/5956
	Δ PK	25 150			5.828	3×10^{-9}	

tive to a preset origin and are permanently recorded on a strip printer. For most of the experiments described below, an AIM-65 microprocessor interface is also utilized to insert the trial-count data on-line into processing routines supported by a TERAK Model 8510 used as a terminal and PDP 11/45 and VAX 750 employing a UNIX operating system programmed in C language. All of the sampling, counting, and display functions can be simply checked by referring them to an internal or external calibrated pulse train generator.

The device also has a manual/automatic option, whereby it will either collect its trial samples only when a panel switch or parallel remote switch is pushed, or it will repeat that process for 50 trials automatically once activated by the switch. The operator thus has the option of triggering each trial or of initiating a repetitive flow of 50 such trials with no further intervention.

The experiments reported here were performed by a single operator, seated in front of the device with the remote initiation switch in hand and the LED count indicators and TERAK terminal display visible. This operator attempted, on instruction or volition, to distort the trial counts either toward higher or lower values. The several options of sampling number, sampling frequency, +/− polarity, and manual/automatic sequencing were variously determined by random instruction, operator preference, or experimental practicality, and recorded before the beginning of each trial. Clearly, the full-matrix of such possibilities could not be explored, and for our first sequence of experiments only 200-sample trials were used, at 100 or 1000 counts/s, all counted in the +/− alternating mode. The automatic/manual and high/low options were more thoroughly tested, in both the volitional and instructed choice modes.

Fifty trials of the 200-sample units comprised a test run, and data from these were processed individually and in many concatenations via a statistics package in the UNIX system developed specifically for this task. Calculated were the mean, standard deviation, range, kurtosis, skew coefficient, z-score, t-score, χ^2 goodness-of-fit with both 8 and 16 degrees of freedom, and the corresponding one-tailed probabilities against chance of the last four measures. Applied to earlier and ongoing baseline data, this analysis confirmed that in undisturbed operation this REG conforms very well to a Gaussian approximation to the appropriate full binary statistics.

The major portion of the results listed below comprised three separate experimental series, extending over fifteen months, labeled REG I, REG II, and REG III, respectively. All other data acquired under slightly less formal conditions of protocol during this period, included for completeness, are grouped under two other series, labeled REG Ia and IIa, respectively. Details of these series protocols, calibration tests, and their individual results are available in the reference [93]. All told, over 25 000 active PK trials were obtained, corresponding to more than 5 000 000 binary events.

Table I summarizes all of the baseline and PK data acquired during these five experimental series. A total of 23 000 base-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

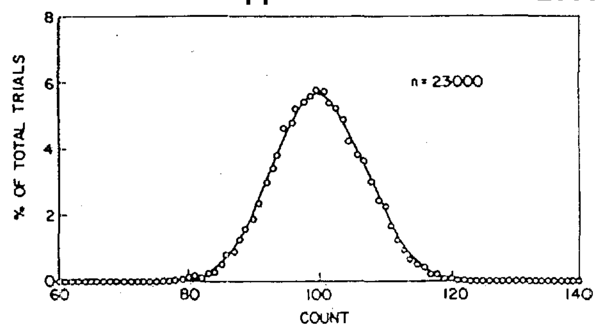


Fig. 11. REG 200-sample baseline data on theory.

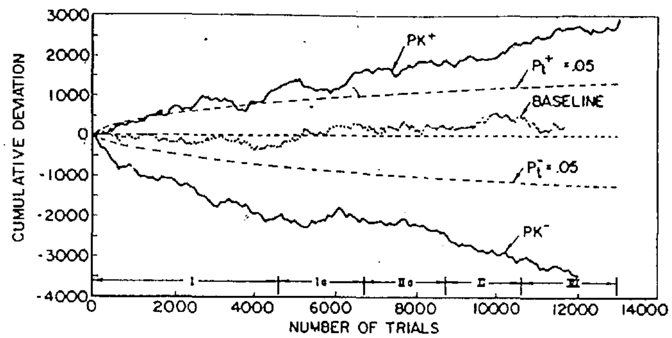


Fig. 14. REG cumulative deviations of PK^+ and PK^- 200-sample data.

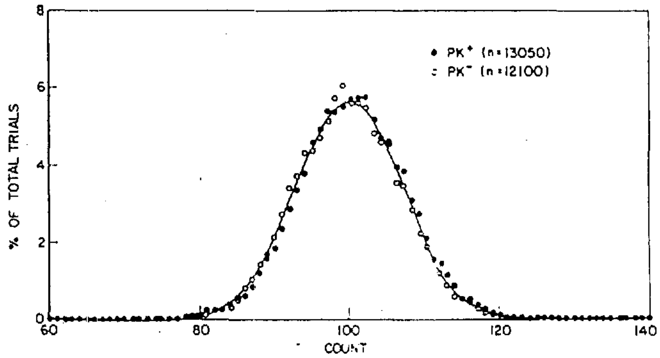


Fig. 12. REG PK^+ and PK^- 200-sample data on theory.

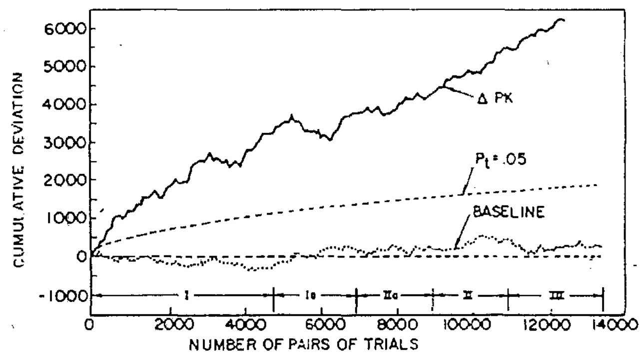


Fig. 15. REG cumulative deviations in direction of effort of all 200-sample data.

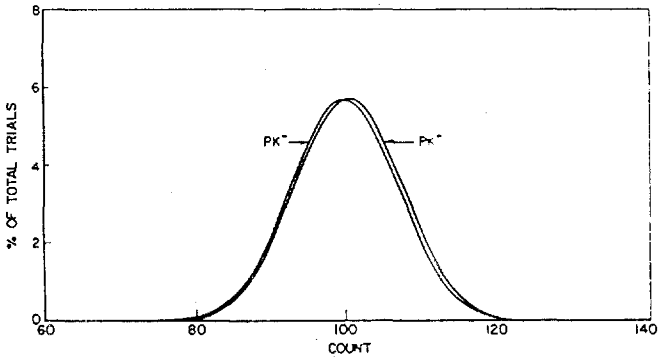


Fig. 13. REG PK^+ and PK^- 200-sample data fitted curves.

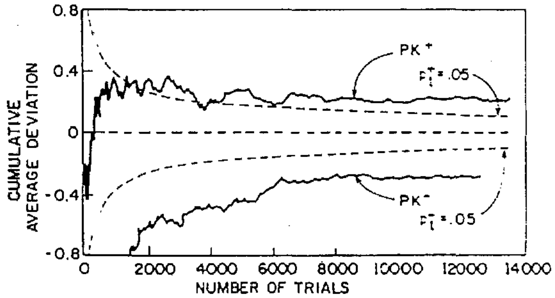


Fig. 16. REG cumulative average deviations of PK^+ and PK^- 200-sample data.

line trials were taken under a variety of conditions before, during, and after the active PK trials. Their overall mean was 100.045, and their standard deviation 6.980, compared with the values of 100.000 and 7.071 predicted by the theoretical Gaussian approximation to the appropriate binary statistical distribution. As shown in Fig. 11, the frequency of count distribution conformed very well with the theoretical curve.

The results of the PK trials are also presented in Table I and in Figs. 12 and 13. Briefly, the 13 050 high-instruction trials, denoted PK^+, yielded a mean of 100.23 and a standard deviation of 6.979; the 12 100 low-instruction trials, denoted PK^-, yielded a mean of 99.704 and a standard deviation of 6.968. The one-tailed probability of chance occurrence of the former, computed from *t*-score, is about 10^{-4}; of the latter, about 2×10^{-6}. The combined probability of the split, i.e., of this total "direction-of-effort" success, denoted ΔPK, is about 3×10^{-9}. (A number of more elaborate statistical measures have been applied to these data; the results are not qualitatively changed thereby.)

As is evident from Figs. 12 and 13, and as is verified by the more detailed statistical tests performed, no significant distortion of the frequency-of-count distributions other than the shifting of the means has occurred. In other words, the observed effect is to shift the total distributions intact, rather than to distort any of their higher moments significantly. This result clearly has felicitous implications for this class of experimentation, since it allows much simpler and faster data collection and analysis than might otherwise have been anticipated.

It is illustrative to exhibit the overall data behavior via graphs of the cumulative deviation of the trial score means versus the accumulated data base. Fig. 14 shows such a representation for the total data pool plotting PK^+, PK^-, and baseline data with reference to cumulative 0.05 confidence levels. Fig. 15 uses a similar representation for compounding the PK^+ and PK^- data in a "direction of effort" cumulative deviation. (Had REG Ia and IIa been excluded from these data, the overall slopes would have been slightly more severe and uniform.)

Alternatively, the cumulative data may be presented in terms of the progressions of the average deviations from the theoretical mean, as shown in Fig. 16, where the stochastic variations

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

of the small data bases are seen to damp out to well-defined terminal values after several thousand trials.

The PK^+ and PK^- effects also manifest themselves in terms of the number of trial means recorded above and below the theoretical value. As displayed in Table I, PK^+ efforts were generally characterized by an excess of trial means above 100.00 and PK^- efforts by an excess below. The total concatenations in these terms are significant at a level of 0.003 for PK^+ and 3×10^{-5} for PK^-.

The ensemble of results acquired in these experiments display certain instructive general features:

1) The importance of accumulating very large data bases when dealing with such marginal phenomena is emphasized by the relative scales of the statistical vagaries and the broader systematic trends in Figs. 14–16. Although the trends are established early in the data collection sequence, unambiguous departures from the accumulated vagaries of chance behavior occur only well into the total 25 000 trial, 5 000 000 bit sequence.

2) Over this large a data base, there arises some quantitative statistical regularity in the PK process, epitomized by the mean slopes of the cumulative deviations in Figs. 14 and 15 and by the terminal values of the average deviations in Fig. 16. Traced back to the elemental binary samples, these values imply directed inversions from chance behavior of about one or one and a half bits in every one thousand or, alternatively, of 0.2 or 0.3 bits per trial.

3) The differences between the somewhat larger values for the PK^- deviation and the lesser values for PK^+ are only marginally significant on this data base, but prevail rather uniformly throughout all the test series. The suspicion that these reflect some subtle bias in the REG itself is not supported by the baseline data, which concatenate to a grand mean very slightly above the theoretical value.

One of the primary goals of such controlled PK studies at this early phase in the understanding of the phenomena is to develop experiments of sufficient yield and replicability that various parametric correlations may be systematically explored, thereby hopefully separating the consequential from the inconsequential factors. The experiments outlined above hold some promise of serviceability for this purpose, but a great deal of data will need be accumulated to establish any such correlations. Four classes of parameters could be considered: those associated with the experimental equipment; those associated with the operator's physiological and emotional characteristics; those associated with the operator's technique; and various environmental factors not directly associated with either. So far we have accumulated only small amounts of data from other operators, and given the general indication #1 above regarding the importance of large data bases, we can make no statement about the generality or peculiarity of our principal operator's performance. Similarly, we have attempted no systematic variation of external environmental factors, and although test times, dates, durations, and laboratory temperature, pressure, and humidity have been routinely recorded, we cannot comment on the importance of this category of parameter.

On the matter of operator technique, it should first be reemphasized that the sole formal difference between the PK^+ and PK^- trials is the specified intention of the operator to influence the device to generate numbers in the assigned direction. No other variation in protocol is permitted, save those subjective differences in psychological attitude the operator chooses to invoke. Although no records of such aspects were kept, this operator, who claimed no special talent for this or any other psychic task, reported that any conscious variations in psychological strategy, such as focus of visual attention, or intensity of concentration or desire, did not appear to have any evident effect on the yield. Similarly, differences in the laboratory ambience, such as the lighting level, background noise, or peripheral presence of other persons, did not seem to influence this operator's performance. When queried about any impressionistic sense of the interaction process, the operator alluded to a "resonance or identification with the system, leading to a loss of self-awareness similar to that experienced in a game, a movie, or some creative occupation." Clearly this class of parameter will be the most difficult to specify and correlate, and we are far from any definition of its mechanisms.

With respect to experimental options on the equipment parameters, we can make very limited explorations with the acquired data base. Briefly, binary correlations of the data for the 100/1000 counting rate option, for the volitional/instructed direction of effort, and for automatic/manual sampling give little indication of importance of such factors in the overall performance. Each category shows clear and significant separation of the means for the PK^+ and PK^- efforts, with little to choose between the t-scores for the various categories. Thus, at least for the data base at hand, the process seems insensitive to these particular experimental parameters.

We have also attempted correlation in terms of the trial-number sequence. With cognizance of the ubiquitous "decline effect" which is reported over a broad range of psychic experimentation, we have prepared an algorithm which cross-concatenates from the data base all scores achieved on the first trials of the experimental run, all achieved on the second trials, etc., up to the fiftieth, and arrays those fifty means in a graphical form. The results show little systematic profile of yield versus trial number. A similar exercise has been performed to cross-concatenate the data by run-number over the various series to search for a decline effect on that larger scale, but again no significant correlation is found within this data base.

The most extensive parametric exploration attempted to date was motivated by the apparently fundamental question implicit in general conclusion 2) above, i.e., whether the magnitude of the observed effect correlates with the total number of bits processed, or with the number of trials. To explore this aspect, the same operator has performed a second ensemble of experimental series totaling 25 000 trials, all consisting of 2000-sample bits rather than 200. As before, various combinations of the automatic/manual and volitional/instructed modes were employed, but to speed data acquisition and reduce the operator's tedium, only the 1000/s counting rate was used. This, coupled with the more elegant data processing capabilities that had evolved over the preceding experiments, allowed this sequence to be completed in less than six months.

The results of this effort, as presented in Table II and Figs. 17 and 18, are curiously ambivalent. As before, there is clear and significant separation of the means of the PK^+ and PK^- efforts, and the baseline is well behaved. As could be anticipated from the larger standard deviation of the 2000-bit data, the cumulative traces display larger statistical fluctuations and require a larger number of trials to settle toward well-defined terminal values. To the quantitative precision allowed by this data base, these terminal values appear not to endorse any simple bit-level hypothesis in that they fail by a factor of 6 or 7 to achieve the one or one and one-half bits per thousand

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE II
SUMMARY OF REG 2000-SAMPLE DATA

Instr.	No. Trials	Mean	Std. Dev.	t-score	P_t	n_+/n_-
B.L.	12 500	1000.016	21.879	0.079	0.468	6157/6088
PK⁺	12 200	1000.380	21.906	1.914	0.028	6092/5897
PK⁻	12 800	999.569	22.005	−2.216	0.013	6218/6351
Δ PK	25 000	—	—	2.920	0.002	

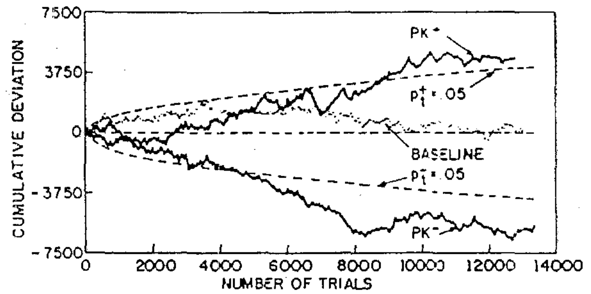


Fig. 17. REG cumulative deviations of PK⁺ and PK⁻ 2000-sample data.

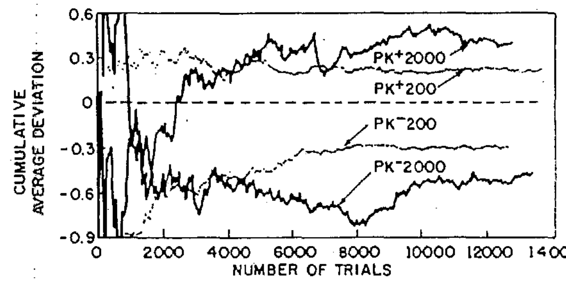


Fig. 18. REG cumulative average deviations of PK⁺ and PK⁻ 2000-sample data compared to 200-sample data.

inversion accomplished in the 200-bit trials. However, the new values are larger on a per-trial basis by a factor of about 1.7, which is not negligible in this context. Again, much more data of this sort will be required to come to grips with this class of correlation.

In addition to continuing study of this sample-size parameter, our next generation of experiments employs a number of other operators to explore the variation of yield with operator type and technique, and a number of alternate noise sources, including pseudo-random sources, in an attempt to localize the effect somewhat and thereby to narrow the range of future experiments and models.

The results outlined above are by no means the only consequential REG data available for contemplation. Of particular interest are a variety of experiments reported by Schmidt, some employing pseudo-random as well as physically random sources, and others using taped source outputs recorded well in advance of their presentation to the operator [87]–[90]. In another approach, May has recently reported an REG study using electronic gear specifically designed to preclude very subtle artifacts that might confound the effects of interest, and includes in his paper a thorough search of the modern REG literature [91].

In addition, considerable research in the parapsychological community has been performed using REG devices as drivers for various forced-choice video games employed in both the PK and clairvoyance modes [151]. Many of these claim significant yields, but rarely are the data-bases sufficiently large to present quantitative trends, or to allow much parameteric correlation.

Regardless of their particular implementation, any potential vulnerability of random electronic noise sources to incidental or intentional distortion by the means under study here would seem to be of some interest to a number of engineering communities, given the proliferate application of such devices in various functional and computational capacities.

Remote Perception

As a second example of contemporary psychic research that has displayed some substantial yield and interlaboratory replicability, we select a topic which has come to be called "remote perception" or "remote viewing." The basic concept of this process is far from new; in the early 16th century, Paracelsus stated it unequivocally:

> Man also possesses a power by which he may see his friends and the circumstances by which they are surrounded, although such persons may be a thousand miles away from him at that time [8].

In its modern form, the experimental protocol requires a "percipient" to describe, by free-response oral or written narrative or drawing, a remote, unknown target location at which is stationed an "agent," with whom there is no normal sensory mode of communication during the course of the experiment. The targets are usually selected by some prescribed random process from a previously prepared pool of targets, which is unknown to any of the active participants. The quality of the perception is assessed by various impressionistic or analytical judging methods described below.

Historically, this experiment has evolved from several generations of free-response clairvoyance and telepathy experiments, which were found to have certain advantages over the more traditional "forced-choice" ESP tasks, such as the Zener card identifications of the early Rhine laboratory [34]–[37], in displaying less tendency for percipient stagnation and "decline-effects" over extended testing, and in maintaining some of the spontaneity of anecdotal clairvoyance experiences. One of the earliest detailed reports of such free-response studies appears as a book by Upton Sinclair entitled *Mental Radio*, which features an equivocal foreword by Albert Einstein [152]. More modern work of this class was performed at the Maimonides Medical Center by Ullman and Krippner in the 1960's, and reported in their book *Dream Telepathy* [64]. From this work emerged the so-called "ganzfeld" or sensory inhibition perception studies of Honorton and many others which propounded the desirability of emotionally stimulating tasks to which the subjects could relate in a personal and spontaneous fashion [66], [67].

The contemporary version of the remote perception protocol was introduced in a sequence of publications by Targ and Puthoff [94], [153]–[156], which prompted a substantial number of attempted replications [157]–[174], and considerable

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Fig. 19. Remote perception target: Picnic area, Feathered Pipe Ranch near Helena, MT; 12:00 N MDT, Sept. 5, 1978.

Percipient transcript: Princeton, NJ; 8:30 A EDT, Sept. 5, 1978:

"Outdoors. . .open landscape–large areas of trees–pines? interspersed with open fields. Single road. High overcast, cool, breezy. (Agent) in dark jacket talking to someone near road–possibly turn-out area or picnic area. Assembly of stones–possibly pylon or marker or wall. Large sign somewhere."

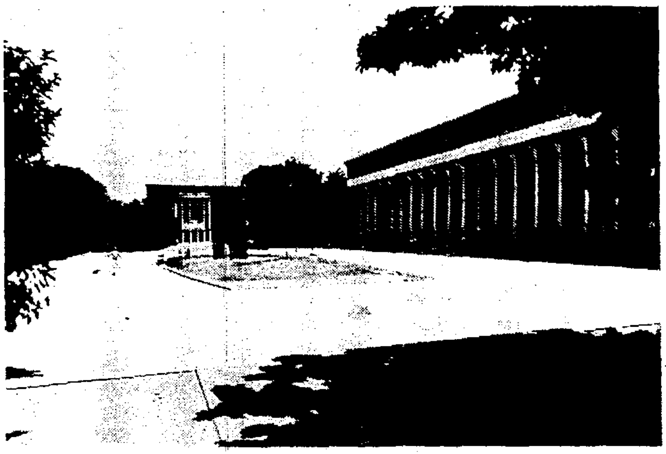


Fig. 20. Remote perception target: Woodrow Wilson School, Princeton, NJ; 2:15 P EDT, Aug. 28, 1980.

Percipient transcript: Princeton, NJ; 12:15 P EDT, Aug. 28, 1980:

"Some kind of courtyard, enclosed by buildings on two sides. Paths or walks around periphery, a statue or monument of some kind in the middle surrounded by grass. Could be a fountain; I have the feeling of water. Trees or tall hedges on one side. Fairly quiet, but some people walking around. Not sure of sound, the idea of a fountain suggests sound of water but I'm not sure I really hear it or not."

critical comment. The most extensive of the replications, conducted by Dunne and Bisaha in the Chicago area over the period 1976 to 1979, comprised 40 formal trials to which were applied 157 independent transcript judgings, 84 of which assigned first-place rank to the proper targets [161], [162].

The type of data which can be acquired in such studies is illustrated in the sequence of Figs. 19–22. In each case is shown a photograph of a particular target, selected by some random process, which was visited by an agent on the date and time indicated. Below each figure is a portion of the corresponding percipient transcript, with the time and location of the perception effort also noted. The examples shown are drawn from a variety of experimental series conducted under somewhat different protocols, but serve to display some of the characteristics which commonly appear in the more successful efforts:

1) The overall ambience of the scene is accurately perceived.
2) Certain details are accurately identified; others are misconstrued or totally ignored.
3) A feature which is impressive to the agent is not necessarily so to the percipient, and vice versa.
4) The composition of the scene may be distorted by errors in scale, relative positions of key objects, or total right-left inversions.
5) The aesthetic aspects, such as colors, general shapes, degree of activity, noise level, climate, and other ambient features tend to be more accurately perceived than more analytical details such as number, size, or relative positions.
6) The perception is not necessarily centered on the defined target, and may even provide accurate information on adjacent areas external to the target, unnoticed by the agent.
7) The fidelity of the perception seems to be independent of the remoteness of the target, up to distances of several thousand miles.
8) The time of the perception effort need not coincide with the time the agent is at the target. Perceptions obtained several hours, or even days, prior to the agent's visit to the target, or even prior to selection of the target, display at least as high a yield as those performed in real time.

The philosophical and practical implications of items 7 and 8 are clearly substantial. If the data are valid, the most parsimonious explications would require access of the percipient's consciousness to other portions of the space-time grid than that in which it is currently immersed, or that it can reach by normal processes of communication or memory. These same items also seriously delimit the potential physical mechanisms for such access.

Rigorous evaluation of the data from experiments such as these is confounded by the psychological components of the process, by the impressionistic nature of the information involved, and by the inevitable subjective biases of all those participating in the experiment. Doubtless the earliest and most primitive assessments were informal *a posteriori* exchanges of impressions about the target between agent and percipient which, although possibly informative and stimulating to them, lacked any quantitative basis and held little scientific credibility. In a somewhat less vulnerable strategy invoked more recently, the percipient, after completing his transcript, visited several possible targets drawn from the pool and attempted to identify the one he perceived, or to rank-order each of them in terms of conformity to his earlier perception. Statistical arguments could then be applied to these ranks to estimate the likelihood that information about the target had been acquired by means other than chance [175]–[177].

In an attempt to separate the possible ESP functioning of the percipient during the visitation and ranking process from his original perception effort, the protocol subsequently evolved to invoke independent judges who were provided copies of the various transcripts and taken to the target sites to perform their preferential rankings. Even in this form, the technique has been criticized for possible sensory cuing of the judges [178] and has tended in turn to be replaced by a protocol wherein the judges perform their ranking on the basis of photographs of the targets, usually taken by the agent at the time of the trial. In one such version, the judges, who have not been involved in any earlier portion of the experiment, are asked to compare a single percipient transcript with agent-generated descriptions and photographs of a number of alternative targets,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

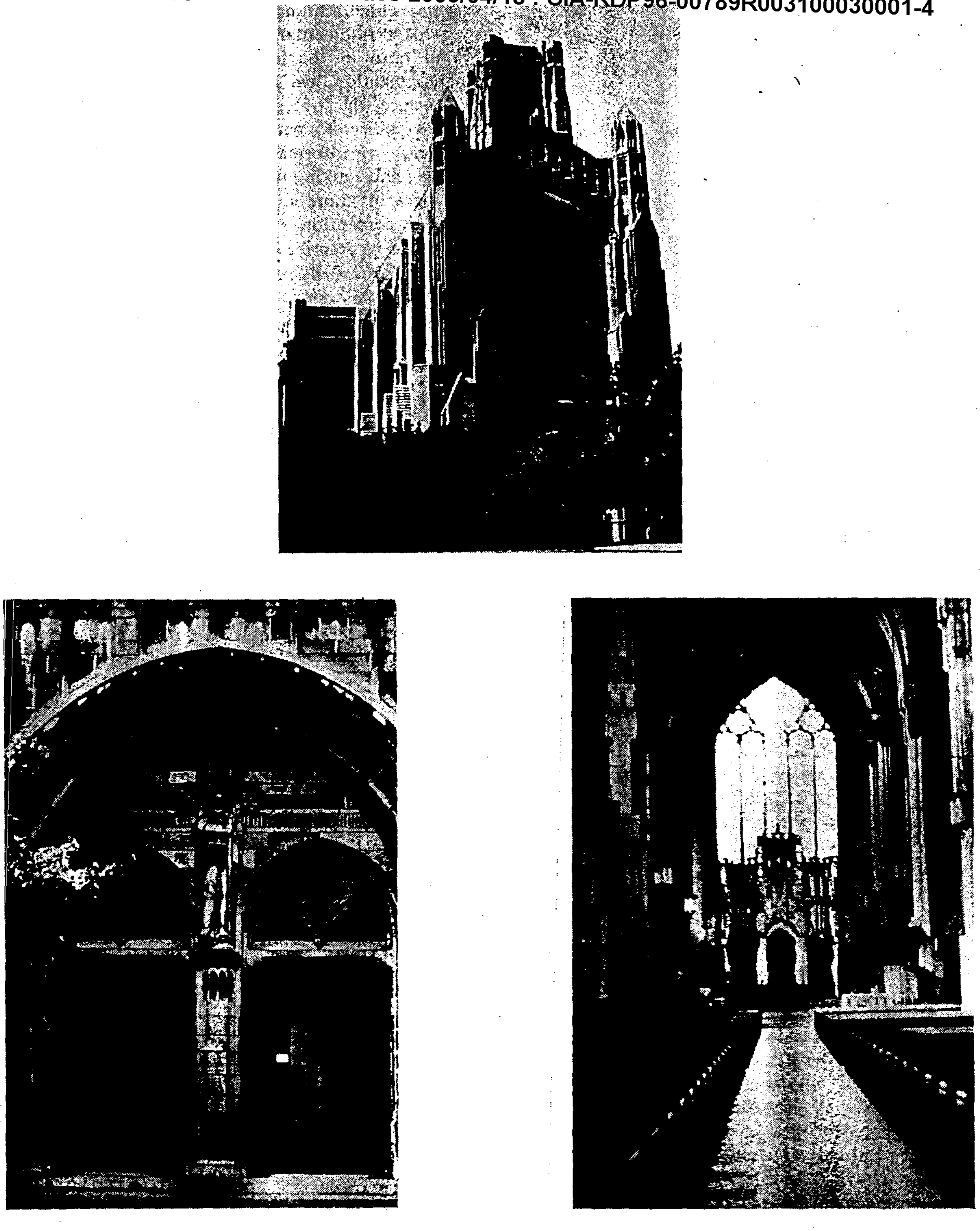


Fig. 21. Remote perception target: Rockefeller Chapel, Chicago, IL; 2:15 P CDT, June 10, 1977.

Percipient transcript: Mundelein College, Chicago, IL; 1:00 P CDT, June 10, 1977:

"I'm seeing a heavy wooden door with a black bolt on it rounded at the top in a dome fashion. I have a feeling of opening the doors and looking in and it's dark inside. My feeling at the moment is that it's a building like a church. And I can see the pews. There is some light but I feel basically a kind of darkness in there and a quietness. I'm seeing little turrets, very elaborate-looking little turrets, a whole series of them like across the entire top of the building and there's a straight line and then up to a triangle. I have a definite image of an angel-type of statue, marble, flowing robes. I see the door again and I see some stairs. I think it's very high. I'm getting some stained-glass windows that are arch shape and they would look to be dark blue. Whatever the architecture of it the ornamentation on the building is quite elaborate and it looks like there's a section on the top with the turrets and then below that there are some other kinds of designs but more linear designs.

"I again have a vision of the doors and then maybe a ledged area or an area of the building that protrudes with some kind of a design and there maybe even be a couple of those before you get to the top part which is either triangular or rounded. There is filigree work, little filigree turrets or something. And within the building there is a sort of a continuation of arches, but possibly they meet columns or something like that, but whatever the decoration is where walls join or separations, it looks like it's arched."

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

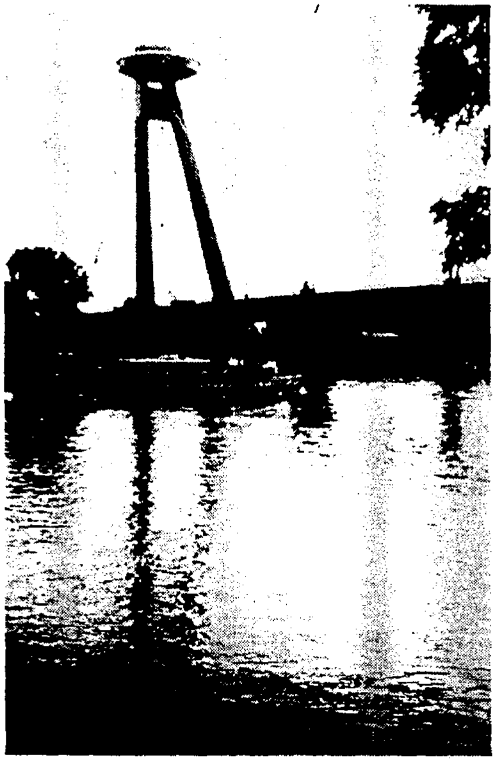


Fig. 22. Remote perception target: Danube River, Bratislava, Czechoslovakia; 3:00 P European Standard Time, Aug. 24, 1976.

Percipient transcript: Minoqua, WI; 8:30 A CDT, Aug. 23, 1976:

"I have the feeling that the agent is somewhere near water. I seem to have the sensation of a very large expanse of water. There might be boats. Several vertical lines, sort of like poles. They're narrow, not heavy. Maybe lamp posts or flag poles. Some kind of circular shape. It's round on its side, like a disc, it's like a round thing flat on the ground, but it seems to have height as well. Maybe with poles. Could possibly come to a point on top. Seeing vertical lines again. Seems to be a strong impression, these vertical lines. Predominant colors seem to be blue and green. Water again. Some very quick impression of a fence, a low fence. The steps seem to go up to some kind of fence. It's a dark fence and it's along like a walk sort of at the top of the steps. The steps sort of lead up to like a path or walkway. Like a boardwalk. And there's a fence along it. There's people walking along it, and there's vertical lines along that walkway."

including the proper target, and to rate or rank-order them by some prescribed criteria. Again, statistical assessment of the significance of the rankings follows.

The bulk of the remote perception data reported in the literature has been evaluated by some form of these independent judging processes, and displays sufficiently high yield to encourage further refinement of the protocol and analysis before attempting categorical judgment on the validity and viability of the phenomena. Beyond minor tightening of the target selection, agent maneuver, and perception acquisition and recording phases of the experiment, the major potential improvements would still seem to pertain to the judging process, which remains potentially vulnerable to subtle cues in the transcripts, to vagaries in the judges' capability, to their subjective biases toward the individual experiments and to the topic as a whole, and to possible psychic input of their own [161], 162].

In an effort to improve the judging process further, our laboratory has explored the applications of various information theoretic methods to the quantification of the data, and an analytic technique has been devised which is based on a limited binary alphabet of target/perception descriptors [179]. While less sensitive to the Gestalt impressionism and symbolic representations which a human judge might capture, this method does provide a rudimentary framework for evaluation of signal-to-noise ratio in the information transfer, and an assessment standardization less dependent on subjective interpretation. In essence, the strategy is to replace impressionistic assessment of the quality of a perception by the identification of specified elements of information therein, after which a mechanical scoring and ranking procedure takes over. In the hope of conveying a bit more substantive flavor of the data acquired in remote perception experiments and the processing thereof, permit us to describe this analysis in a little detail.

The heart of the method is the establishment of a code, or alphabet, of simple descriptive queries which may be addressed to all targets and all perceptions, responses to which serve to distinguish them and to permit quantification of the information acquired in the perception process. In one version, these "descriptors," thirty in number, are posed in binary form and range over a spectrum from quite factual discriminations, e.g., whether the scene is indoors or outdoors, whether trees are present, or whether there are automobiles, to much more impressionistic aspects, such as whether the ambience is noisy or quiet, confined or expansive, hectic or tranquil. The particular ensemble of descriptors has evolved in part through personal experience and intuition, and in part through trial-and-error application to various pilot data. The goal has been a balanced alphabet whose elements are a) relatively unambiguous; b) commonly perceived by a broad selection of percipients; c) individually instructive in defining the scene; d) complementary to one another; and e) sufficient in number to permit reasonable synthesis of the scene, but not so numerous to burden the data collection or computation excessively.

Given this descriptor alphabet, each target in the pool is then represented in terms of 30 binary bits, corresponding to the appropriate YES/NO responses to the queries. This encoding is normally performed by the agent at the time of target visitation, although reference may be made to the target selector's judgment or to photographs of the target for verification. Each perception is similarly rendered into a corresponding sequence of binary digits, but only after the percipient has been allowed to form a free-response impression of the target. Various scoring recipes are then invoked for quantitative comparison of the perceptions with the targets, using for computation the UNIX operating system of a PDP 11/45 or VAX 750.

The simplest recipe merely counts the number of correct responses to the 30 descriptors, i.e., the positive correlations between the target and descriptor matrices. This does not normally provide a particularly accurate index of the quality of the individual perceptions, since the *a priori* probabilities of the various descriptors are widely different. For example, a given pool may have more outdoor than indoor targets, and hence a correct identification of an indoor context should be given higher credit than identification of an outdoor context. To facilitate such weighting, a step is included in the computational program to provide the *a priori* probabilities of all descriptors in the prevailing target pool, on the basis of which more elaborate scoring recipes may be invoked.

Since the various targets have substantially different characteristics and hence different capacities for achieved scores, a variety of normalization procedures also have been developed, using as denominators the total number of descriptors, the perfect score, i.e., the score that would be achieved for a given target if all descriptors were identified correctly, and various "chance" scores for the target, defined by some random or

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

PRECOGNITIVE REMOTE PERCEPTION 24 × 24 "CHICAGO" SERIES
PROPER TARGET RANKS

Perception No.	Scoring Method[a] A	B	C	D	E	Avg.	Avg./24	Human Judges Mean Rank[b]
1	3/4	7	7	9	2	5.9	0.25	2.7/8 = 0.34
2	1/2	2	4	1	1	1.9	0.08	1.0/8 = 0.13
3	1	1	1	1	1	1.0	0.04	1.5/8 = 0.19
4	2/8	7/2	2/2	4/2	2/2	4.5	0.19	2.7/8 = 0.34
5	7/3	9	11	4	8	8.0	0.33	1.7/8 = 0.21
6	9/5	12/2	7/2	16/2	7/2	11.0	0.46	3.5/8 = 0.44
7	13/2	11	14	10	13	12.3	0.51	2.3/8 = 0.29
8	20/3	22	20	14	19	19.2	0.80	1.8/8 = 0.23
9	4/2	4	8	1	5	4.5	0.19	2.6/7 = 0.37
10	10/7	13	13	9	5	10.6	0.44	1.4/7 = 0.20
11	9/4	9	11	6	12	9.7	0.40	3.6/7 = 0.51
12	1/3	2	2	2	5	2.6	0.11	1.8/7 = 0.26
13	1/3	1	3	1	3	2.0	0.08	2.2/7 = 0.31
14	1	2	2	1	1	1.4	0.06	1.4/7 = 0.20
15	1/2	2	2	1	1	1.5	0.06	1.0/7 = 0.14
16	2	2	2	1	2	1.8	0.08	1.0/10 = 0.10
17	1/4	9	1	9	1	4.5	0.19	1.0/5 = 0.20
18	2/2	10	5	1	2	4.1	0.17	1.0/5 = 0.20
19	14/3	14	14	19	17	15.8	0.66	unjudged
20	7/6	11	11	8	10	9.9	0.41	5.0/6 = 0.83
21	1/2	2	3	4	1	2.3	0.10	2.0/6 = 0.33
22	5/6	7	12	2	7	7.1	0.30	3.0/6 = 0.50
23	16/4	23	11	15	9	15.1	0.63	3.0/6 = 0.50
24	3	3	4	3	4	3.4	0.14	2.0/6 = 0.33
Mean	6.73	7.75	7.13	5.96	5.79	6.67	0.28	0.31

[a] Computed rank/number of ties for that rank.
[b] Assigned rank/number of possible ranks.

arbitrary process of descriptor response. A "selective" scoring/normalization process has also been applied which effectively allows the percipient to reject any descriptor on which he feels unqualified to comment, and thence to be scored only on the reduced descriptor set.

The statistical significances of these various normalized perception scores are assessed by a collective ranking process reminiscent of the traditional human judging techniques, but having the advantages that the ranking proceeds on a much more standardized and analytical basis, and that many more alternative targets can be ranked by the machine than by a human judge. Specifically, the program scores each transcript not only against its proper target, but against every other target in the pool, and then ranks these targets in order of descending score and specifies the rank of the match with the proper target. This process is repeated for every scoring method, and the results displayed in corresponding matrix arrays.

Table III displays typical results of these analytical ranking procedures as applied to a group of 24 perceptions of 24 targets in the Chicago area. Tabulated are the ranks of the proper targets compared with all other targets for each of the perception efforts, as computed by five of the scoring methods we have found to be most instructive, namely, A) number of correct descriptors/total number of descriptors; B) weighted full descriptor score/perfect score; C) weighted full descriptor score/number of descriptors; D) weighted selective descriptor score/perfect score; and E) weighted selective descriptor score/chance score. Also included in the table are the mean ranks assigned by independent human judges subjectively comparing these perceptions with a much smaller number of alternative targets. Although the bases of comparison are quite different, it appears that in the majority of these cases the analytical and impressionistic evaluations concur at least roughly in their estimate of the quality of the perceptions, particularly for those which consistently obtain low rank assignments. If the analytical computation is carried through using as target pool only those alternative targets available to the human judges, the agreement in mean ranks is found to be somewhat closer, perhaps fortuitously so, given the categorically different bases of assessment implicit in the two methods.

To this analytically scored and ranked data it is possible to apply a variety of statistical assessments of widely ranging sophistication and complexity. Consistent with the rather broad mesh of the descriptor code and the elementary scoring recipes invoked in this version of the concept, we confine ourselves to correspondingly simple statistical measures which provide at least semi-quantitative indication of the yield beyond chance. Specifically, we address only the distribution of *proper* target ranks achieved in the series of perceptions, such as summarized in columns 2–6 of Table III. Using the common z-score method for a discrete distribution, the probability of achieving the mean rank of any of these columns by chance may be directly computed. Table IV displays the results of such calculation for the same 24 × 24 "Chicago" series. Note that, whereas all of the methods suggest significant departures of the computed mean ranks from chance, there is relatively little disparity among them, indicating that the specific method of scoring and normalization is not a sensitive element in the overall evaluation of the perception series.

The departure of the shape of the proper target rank distribution from chance is also displayed in Table IV in terms of the number of perceptions achieving first-place ranks, the number

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE IV
PRECOGNITIVE REMOTE PERCEPTION 24 × 24 "CHICAGO" SERIES

Method[a]	Mean Rank	No. 1st (2nd) Ranks	No. Ranks Below Mean	z	P_z	$\chi^2(4)$	P_χ
Chance expectation	12.5	1.0 (1.0)	12				
A	6.73	4.4 (4.0)	19	-4.08	2×10^{-5}	18.5	<0.005
B	7.75	2.0 (6.0)	20	-3.36	4×10^{-4}	11.8	<0.01
C	7.13	2.0 (4.5)	20	-3.80	7×10^{-5}	19.6	<0.001
D	5.96	8.0 (2.0)	20	-4.63	2×10^{-6}	24.0	<0.001
E	5.79	6.0 (3.5)	21	-4.75	1×10^{-6}	27.6	<0.001

[a] As described in text:
A) Number of correct descriptors/total number of descriptors.
B) Weighted full descriptor score/perfect score.
C) Weighted full descriptor score/number of descriptors.
D) Weighted selective descriptor score/perfect score.
E) Weighted selective descriptor score/chance score.

TABLE V
PRECOGNITIVE REMOTE PERCEPTION 6 × 24 "OCEAN" SERIES
SCORE SUMMARY

Method[a]	Mean Rank	No. 1st (2nd) Ranks	No. Ranks Below Mean	z	P_z
Chance expectation	12.5	0.25 (0.25)	3		
A	4.5	1.0 (1.1)	6	-2.83	0.002
B	4.2	2.0 (0.0)	6	-2.95	0.002
C	5.0	1.0 (1.0)	6	-2.65	0.004
D	5.2	1.0 (0.5)	6	-2.59	0.005
E	5.3	1.0 (1.0)	6	-2.54	0.006

[a] As described in text:
A) Number of correct descriptors/total number of descriptors.
B) Weighted full descriptor score/perfect score.
C) Weighted full descriptor score/number of descriptors.
D) Weighted selective descriptor score/perfect score.
E) Weighted selective descriptor score/chance score.

ranked better than the chance mean, and a simple χ^2 test with its associated probability.

The method is not restricted to square arrays, i.e., to equal numbers of targets and perceptions in one-to-one correspondence. Table V displays the results of a recent "Ocean" series in which six perceptions were ranked against 24 targets in the pool from which the six actual targets were drawn. (The percipient for this series was at the time sailing alone across the North Atlantic Ocean.)

Because of its particularly severe protocol, we also include as illustration the results of a "European" series judged by this method. In this series, the agent was traveling in eastern Europe, and on five successive days, at 3:00 P.M. local time, utilized whatever location he happened to occupy as the target. The perceptions of these targets were recorded by a single percipient in northern Wisconsin at approximately 8:30 A.M. local time of the preceding day, i.e., each perception was roughly 24 h precognitive. (Fig. 22 depicts a target and perception from this series.) Table VI displays the analytically judged results of this series, and compares them with the results of previous human judging. Although the data base is small, the consistency of yield is striking.

To obviate the possibility that this method of analysis may somehow process even random inputs to apparently significant scores, artificial target data matrices and artificial perception data matrices have been constructed from the output of a REG, and the computational schemes applied to various combinations of these with each other and with true data. The pattern of results conforms to chance expectations. An alternative form of control is provided by application of the method to discernibly unsuccessful test series, which yield appropriately insignificant results.

The project outlined in the last few paragraphs is described in greater detail, with many other experimental examples, both successful and unsuccessful, in reference [179]. We have since developed the capability of employing ternary rather than binary responses to the descriptors, in order to convey more shaded information about the aspects queried. A given feature may thereby be specified as a) present and dominant; b) present but secondary; or c) absent. Or alternatively, the feature may be described as a) definitely present; b) ambiguous or unspecified; or c) definitely absent. While these approaches clearly provide more specific target and perception data, the scoring thereof becomes more complex, especially in the definition of certain of the normalization denominators.

We have also been exploring a modification of this analytical judging procedure which would bypass the ranking steps altogether and move directly to compute individual statistical scores for each transcript. The key to this variation is the definition and utilization of a generalized target pool composed of ove

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE VI
PRECOGNITIVE REMOTE PERCEPTION 5 × 5 "EUROPEAN" SERIES
SCORE SUMMARY

Method[a]	Mean Rank	No. 1st (2nd) Ranks	No. Ranks Below Mean	z	P_z
Chance expectation	3.0	1.0 (1.0)	3		
A	1.2	4.3 (0.3)	5	-2.85	0.002
B	1.4	4.0 (0.0)	5	-2.53	0.006
C	1.4	3.0 (2.0)	5	-2.53	0.005
D	1.8	4.0 (0.0)	4	-1.90	0.029
E	1.6	3.0 (1.0)	5	-2.21	0.013
Seven human judges (avg.)	1.9	2.1 (1.6)	4.6		

[a]As described in text:
A) Number of correct descriptors/total number of descriptors.
B) Weighted full descriptor score/perfect score.
C) Weighted full descriptor score/number of descriptors.
D) Weighted selective descriptor score/perfect score.
E) Weighted selective descriptor score/chance score.

200 targets, local, national, and international, assembled for a broad range of earlier and ongoing experimental series, from which correspondingly generalized *a priori* descriptor probabilities may be calculated. Evaluation of the transcript scores on the basis of these generalized probabilities, rather than those calculated for the specific series target pools, has been found to alter only slightly the relative ranks of the perceptions determined by any of the five methods used above and provides the desired common basis for evaluating the individual statistical significance of those scores.

To pursue this, the program next assembles a set of empirical chance distribution functions, one for each scoring method, by concatenating all mismatched perception scores assigned by that method, i.e., all off-diagonal elements of that perception-target matrix. With reference to Gaussian fits to these empirical chance distributions, the proper perception-target scores can then be assigned z-values and corresponding probabilities against "chance." As one example of this approach, Fig. 23 shows the empirical distribution of scores compounded from some 1400 mismatched targets and transcripts by method E, on which are superimposed the scores achieved by the 24 proper target perceptions of the "Chicago" series. The corresponding z-values and probabilities against chance are listed in Table VII.

These individual significance values can subsequently be compounded into an overall significance level for the entire series by various standard procedures [180]. The latter result should agree with that derived from the original ranking method, to the order of approximation implicit in each form of this analysis. In this case, the series value is about 10^{-8}, compared to 10^{-6} for the ranking method.

The specific results shown in Tables III–VII represent some of the most successful data we have acquired; many less successful examples could also be displayed. To summarize our total experience with over two hundred remote perception efforts, all performed with volunteer percipients claiming no special abilities, we might note that the data tend to fall into one of four categories, in roughly comparable quantity:

1) the target is accurately represented in detail and composition;
2) particular features of the target are accurately perceived, but the context is incorrect;
3) the ambience of the target is perceived, but the details are inaccurate;

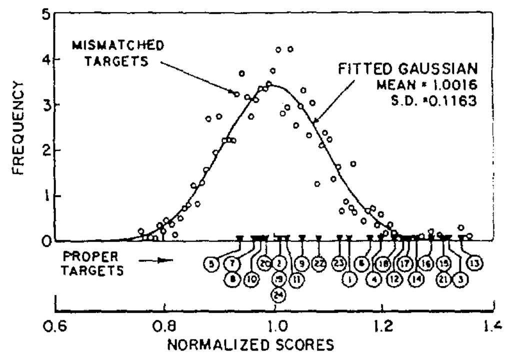


Fig. 23. Remote perception empirical chance distribution (method E).

4) the perception seems unrelated to the target in context or detail.

Survey of the pertinent literature indicates a comparable pattern of yield across the experience of others involved in similar experiments.

To date there has been little progress in correlating the degree of success of such efforts with the prevailing experimental conditions or with the personality traits or attitudes of the participants. A certain body of lore has compounded from the testimony of the more successful percipients, such as the desirability of a personal rapport between the agent and percipient, the value of a lighthearted attitude, the importance of excluding any associative or constructive logic, etc., but much of this is still too vague and inconsistent to provide any basis for experimental refinement or theoretical modeling. At present, the only fair statement would be that empirical evidence for this class of phenomena continues to accumulate, but with frustrating irregularity and little basic comprehension. Notwithstanding, the present and potential applications of the process in a variety of arenas, combined with the relatively simple and inexpensive nature of the experiment, keep this type of study active.

THEORETICAL CONCEPTS

No remotely satisfactory physical theory of psychic phenomena yet exists. Indeed, next to the evasiveness of the effects under controlled experimentation, the second greatest frustra-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE VII

PRECOGNITIVE REMOTE PERCEPTION INDIVIDUAL STATISTICAL SCORES (METHOD E) 24 × 24 "CHICAGO" SERIES

Perception No.	z-score	P_z
1	1.210	0.113
2	0.089	0.465
3	2.737	0.003
4	1.700	0.045
5	-0.553	0.709
6	1.533	0.063
7	-0.315	0.623
8	-0.315	0.623
9	0.451	0.326
10	-0.197	0.578
11	0.191	0.424
12	2.037	0.021
13	2.948	0.002
14	2.244	0.012
15	2.661	0.004
16	2.478	0.007
17	2.180	0.017
18	1.904	0.028
19	0.083	0.467
20	-0.152	0.560
21	2.679	0.004
22	0.701	0.241
23	1.057	0.145
24	0.113	0.455

tion in the study of psychic processes has been the absence of viable theoretical models with which to begin the traditional dialogue between theory and experiment on which all scientific progress eventually depends. This may, of course, be indicative of an illegitimacy of the phenomena themselves, or at least of an evanescence that fundamentally precludes any analytical representation. On the other hand, it may be an indication that modern physical theory, elaborate and sophisticated as it is, has not yet evolved to a stage where it can properly acknowledge and deal with the role of consciousness in the physical world [181], and that this should be one of its new frontiers. Before pouncing toward either extreme, it may be worthwhile to play through some of the more canonical attempts that have been made to deal with this domain, both from a formalistic and philosophical point of view.

Efforts toward theoretical explication of psychic phenomena over the past century have proceeded from various levels of presumption as to the fundamentality of the effects observed. Some have insisted that the effects are totally illusory, i.e., artifacts of poor experimentation or data processing, or that they are the chance results of random processes. Others have assigned the effects to known physical and physiological mechanisms associated with, but not deliberately precipitated by the participants, e.g., electromagnetic radiation from brain circuitry or intercardial potentials, or heat transfer, vibration, or aerochemical changes in the experimental environment caused by human presence or exertion. More ambitious efforts have contended that no such conservative models can suffice, and that it will be necessary to identify new forms of energy or information transfer to retain the established physical formalisms, or possibly even necessary to expand the physical laws themselves, as was required in the generalization from classical mechanics to quantum mechanics or to special and general relativity, with the present forms becoming subsumed under more comprehensive statements. Still others have concluded that the scientific paradigm in general is inadequate and that basic revision in the representation of the process of conscious observation of physical events will be required.

Beyond the uncertainty as to the fundamentality of the model required, these modeling efforts have labored under the confusion of whether the purported phenomena are most basically psychological, physiological, physical, or some inextricable combination thereof, and hence which class of concepts should dominate the model and which should be secondary. Virtually all permutations have been explored to some degree, under labels of "psycho-physiological," "biophysical," "psycho-physical," etc., but none of these can claim more than suggestive analogies or philosophical stimulation. This author is unqualified to assess any models based in the psychological or physiological domains, other than to note that those most frequently discussed tend to acknowledge the role of random processes, information ordering, and statistical, rather than directly causal mechanisms [182]–[187], all of which have their counterparts in several of the more physically oriented modes which have been proposed. Confining ourselves to such physical theories, the history of credible analytic effort is conveniently short and, in my view, more instructive in its philosophical than in its functional characteristics.

Electromagnetic and Geophysical Models

For perhaps naive reasons, the earliest physical models tended to presume wavelike propagation of psychic effects, usually in the electromagnetic modes [188]. Doubtless the concurrent emergence of radio technology with its revolutionary wireless capabilities influenced the concepts and nomenclature of these versions, and frequent reference to psychic "transmitters," "receivers," "antennas," "tuning," and "static" are found in them. Upton Sinclair's book, *Mental Radio*, mentioned earlier, is one example of such an analogy [152].

More modern models of this genre, appearing predominantly in the Russian literature [43], [50], [100], [189], [190], have focused on very low frequency bands, of the order of 10 Hz,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

characteristic of various physiological frequencies, especially the brain wave spectra. Some variations of these have invoked modulation of the earth's magnetic field or of the electrostatic fields of the atmosphere. Wave models involving other than electromagnetic environmental media have also been proposed, such as infrasonic atmospheric waves, geoseismic waves, and barometric fluctuations, possibly stimulated by contemporary attention to the inexplicable homing capabilities of birds, fish, and animals, and the hive or swarm behavior of certain insects [101], [102], [132].

As suggested by the pre-occupation with screen-rooms in the early days of the electromagnetic concept, and by more recent long-distance remote perception experiments, some of the obvious validation/disqualification tests examining attenuation, diffraction, interference, and polarization effects have indeed been attempted, but the very large dimensions involved for these wavelengths leave the studies less than conclusive. Some contemporary workers retain support for this category of model, claiming that the necessary electromagnetic signal levels required for influence on the brain circuitry are so small that no coarse-grained tests can properly discriminate against them. More problematic, in my own view, are the absence of any demonstrated velocity of propagation of psychic effects and the large body of precognition evidence which, if accepted, cannot be accommodated by any reasonable advanced wave characteristics.

However, setting aside for the moment the quantitative functional difficulties with the electromagnetic approach to psychic phenomena, certain broader philosophical analogies in the conceptual and experimental aspects of the two topics may be worth noting. In electromagnetism, beyond the bald empiricism of the definition of the fields themselves as representations of "action-at-a-distance," there are many features which to a naive or primitive observer would, and indeed historically did, appear as "paranormal," or at least anomalous: the noncolinear aspects of Ampere's and Biot-Savart's laws and of the Faraday effect; induction effects and switching transients; the Maxwell displacement current; the propagation of waves in a vacuum, with a specified finite velocity—each of these in some sense digressed from contemporary "normal" experience, was conceptually difficult in its time, and required certain leaps of empiricism to get on with the formalism.

To the extent that we just now hold a similarly naive and primitive view of psychic phenomena, it may be necessary to tolerate similarly empirical representations until a more comprehensive model can knit itself into a more fundamental representation. For this purpose, there may be some merit in looking to just such electromagnetic effects for analogies. As one example, the pervasive "decline effect" in psychic experimentation, wherein the performance on psychic tasks is widely reported to be highest immediately after initiation, to decrease over protracted testing periods, and then to improve again just before termination, is somewhat reminiscent of the induced signatures of certain electromagnetic processes. The decline effect has been commonly ascribed to a psychological boredom or reduction in the emotional intensity of the operator performing the task; just possibly it may be a more fundamental characteristic of the phenomenological domain.

In a similar vein, many psychic effects are reported to be intrinsically transient, e.g., the "fleeting impression"; the "sudden vision"; the "unexpected effect." One of the favored techniques of some remote perception percipients is to "sweep through" their image of the target repeatedly until it is clarified. Many PK effects are reported to be achieved just at the first effort or immediately after the effort is terminated. Such "beginner's luck" evidence might be construed to indicate that psychic processes are invalidated by prolonged and careful examination; alternatively, it may be a hint that they are inherently "inductive," in the electromagnetic sense, i.e, that they are intrinsically unsteady phenomena wherein the time derivatives influence the magnitudes of the effects.

Yet another characteristic that psychic effects share with certain electromagnetic processes is their tendency to be indirect, tangential, or peripheral: direct effort on one PK task fails, but a secondary effect is noted on another component or device; central elements of a remote perception target are ignored by the percipient, but minor or peripheral aspects are identified with precision; the pattern of physical disturbance in a poltergeist event is reported to be vortical rather than radial, all of which call to mind cross-product and vector curl effects in the electromagnetic domain.

The point in suggesting such analogies is not to endorse direct physical correspondence between electromagnetic and psychic processes, but rather to speculate whether the human mind may tend to perceive and assess phenomena in the two domains in certain similar fundamental ways.

Entropy and the Random Process

A second, more recently opened class of psychic model addresses the interaction of consciousness with natural random processes [95], [187], [191]. A common aspect of the established physical formalisms of kinetic theory, thermodynamics, statistical mechanics, and information theory is the role of randomicity as the reference plane for information and energy exchange. By whatever representation, the second law of thermodynamics, expressing the tendency of isolated physical systems to drive irreversibly toward configurations of minimum order and information content, stands starkly asymmetric in the time coordinate, thereby raising profound philosophical issues in virtually every domain from biophysics to cosmology.

Some of the most controlled and replicable experiments in PK, such as the REG studies outlined earlier, could be construed to challenge the second law, or at least to suggest modifications of the concept of an isolated physical system. Namely, under the circumstances of those experiments, human consciousness could be postulated to be inserting order, albeit to a small degree, into a random physical process.

This possibility can be extended conceptually to the anomalous acquisition of information in remote perception experiments, to psychic healing, and to animal and plant PK, but to my knowledge, no attempts at formulation have yet specified any details of the ordering capability, e.g., its physical or physiological source, its propagation modes, or its manner of interaction. Pending these, one can again only proceed with high empiricism to attempt to represent the observed correlations in a useful fashion, a strategy which has sustained many other observational fields in their primitive phases. Notwithstanding, this class of psychic model poses a profound question: The long-accepted essence of consciousness is its ability to extract information from its environment; may the reverse also be possible? May consciousness have the ability to insert information into its environment?

One extreme variation of this model escalates the question even one step further, to ask whether it is possibly an indigenous property of extremely elaborate and complex systems that they may embody inherent functional consciousness of

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

their own—that somehow out of their very complexity, and interlockings of their systems, they derive not only abilities to learn, to reproduce themselves, to adapt to their environment, but also to exert an entropy reversing form of "consciousness" on themselves [192], [193].

Hyperspace Representations

A few attempts have been made to represent paranormal effects by re-casting the basic laws of physics in more than the four coordinates of normal human experience and applying the consequent new terms to the representation of paranormal effects [104], [105], [194]. For example, one such approach adds an imaginary component to each of the spatial and temporal coordinates, i.e., invokes complex space and time, somewhat in the spirit of ac circuit theory or exponential wave mechanics [195]. The imaginary components and their "cross-talk" with the traditional real ones thereby permit representation of anomalous effects within the framework of established physical laws. To date, no convincing fundamental definition of these new dimensions has been offered, thereby reducing such models to an heuristic or empirical level. In this sense, the approach bears some similarity to the "hidden variable" aspect of quantum mechanics, mentioned below.

One may speculate that if such models are to address the interaction of conscioueness with physical process, the requisite new coordinates or components must relate to, or in some way define or localize, the processes of consciousness. In other words, to the normal "hard" coordinates, whereby events are conventionally specified in the physical world, it may be necessary to add certain "soft" coordinates to specify the conscious processes by which those events are perceived and possibly influenced. Clearly, the coordinate frame in which one chooses to observe and represent any physical process is intimately linked to the perception of that process. Two common examples would be the perceptions of kinematic and dynamic effects in a rotating frame of reference, such as a merry-go-round or an orbiting spacecraft, or the even more bizarre appearances of physical processes in rapidly accelerating frames, à la general relativity.

To compound this interdependence of perception and reference frame by including "consciousness coordinates" in the specification of the latter is an awesome proposition, but an intriguing one. And the consequences need not be restricted to the mechanical behavior of physical systems, but may also influence their perceived substance. The noted British astronomer Sir Arthur Eddington, some fifty years ago, presaged this concept most boldly and heretically [196], [197], [203]:

> The whole of those laws of nature. . .have their origin, not in any special mechanisms of nature, but in the workings of the mind. . . .
>
> All through the physical world runs that unknown content which must surely be the stuff of our consciousness. . . .
>
> Where science has progressed the farthest, the mind has regained from nature that which the mind has put into nature. . . .
>
> We may look forward with undiminished enthusiasm to learning in the coming years what lies in the atomic nucleus—even though we suspect that it is hidden there by ourselves. . . .
>
> The stuff of the world in mind-stuff.

Transform Models

Another rather extreme approach proposed recently has come to be referred to as the "holographic" or "transform" model [198]–[200]. Essentially what is suggested here, as I understand it, is that the information of the universe is arrayed, not in terms of position and time as we have come to perceive it, but rather as frequency and amplitude information, and that the human consciousness essentially performs "Fourier transforms" on this to order that information into the more familiar form. In the sense that the space and time coordinates are thereby downgraded from the fundamental coordinates of experience to useful ordering parameters, one could interpret such models to imply that consciousness may, by this mechanism, access any portion of space and time to acquire information, and then interpret it in some characteristic form.

The physicist David Bohm, in conjunction with the psychologist Karl Pribram, has elaborated such concepts to a considerable degree of generality, proposing a so-called "implicate order" or "enfolded order" of fundamental reality from which the more familiar "explicate order," i.e., the commonly manifest perceptions are assembled in accordance with the prevailing circumstances of their observation [201], [202].

Quantum Mechanical Models

Probably the most exercised category of contemporary model attempts to apply the concepts and formalisms of quantum mechanics to represent some of the paranormal effects presented in the psychic domain. Of all the forms of physical analysis, quantum mechanics invokes the largest array of empirical postulates that are at variance with conventional rationality, and yields in their implementation a corresponding array of results which contradict common impressions of reality. The quantization process itself, which limits measurable properties to discrete values; the representation of particulate systems by wave functions; the role of observation in collapsing the wave functions to a single state vector; the uncertainty principle; the exclusion principle; the indistinguishability principle; and most drastic of all, the commitment to totally probabilistic mechanical behaviors—all, in some sense concede a degree of paradox in human perception of physical processes. The familiar conundrums of "Schrödinger's Cat," "Wigner's Friend," or the "Einstein–Podolsky–Rosen Paradox," all suggest that the laws of quantum mechanics are not so much statements of fundamental physical reality, as of our ability to acquire information about that reality. Quantum mechanics, in other words, does not so much describe the state of a physical system as it describes our *knowledge* of the state of that system.

It is somewhat in this spirit that a number of authors have aspired to model psychic process in quantum mechanical terms. Some have attempted to invoke the so-called "hidden" or unused variables of the formalism to involve conscious process more explicitly in the behavior of physical systems [107], [204]–[208]. Others have endeavored to draw analogies between the synaptic processes in the brain and quantum mechanical "tunneling" [209].

Given the primitive state of the phenomenological data base, much of this effort may be prematurely elaborate and complex. Our own approach to quantum mechanical modeling has been far more superficial and generic, attempting only to explore possible analogies between the paradoxical consequences of the formalism, and the paranormal evidence of certain psychic experiments. Again with the indulgence of the reader, we might sketch a bit of this argument, for the purpose of illustrating this class of approach with the example closest to hand.

One conventional interpretation of the application of quantum mechanical formalism to the observable behavior of physical systems is to associate appropriate mathematical operators with a corresponding measurement process. When applied to

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the prevailing wave function of the system, these operators call out the observable values of the property in question as eigenvalues of an equation of the form

$$M\psi_i = m_i\psi_i$$

where M is the measurement operator, m_i are the observable values of the measured property, and ψ_i the corresponding eigenstates of the wave function. Our approach is to generalize this representation to include conscious systems as well as conventional physical ones, and to allow the measurement operator concept to include specification of psychological as well as physical properties.

Thus, we denote a particular individual consciousness by a "state function" ψ_i, and represent a situation to which it is exposed by an operator S. Application of the situation operator to the consciousness wave function then yields the possible psychological responses, s_i, as eigenvalues:

$$S\psi_i = s_i\psi_i.$$

We then invoke certain aspects of quantum mechanical interaction theory to develop the capacity for "paranormal" behavior of both physical and conscious systems. For example, in the traditional theory of the covalent chemical bond between two hydrogen atoms, one constructs from the separate atomic functions, ψ^a and ψ^b, using arguments of symmetry and indistinguishability, a composite molecular wave function ψ^{ab}, which yields expectation values for the molecular energy levels substantially different from simple linear superposition of the atomic energy eigenvalues, i.e.,

$$e_i^{ab} = e_i^a + e_i^b + \Delta e_i^{ab}$$

where e_i^a and e_i^b are the energy eigenvalues of the atomic systems, and Δe_i^{ab} embodies an "exchange energy" term which is classically inexplicable, but devolves formally from the postulate that the electrons are indistinguishable in the bonded configuration. Stated more bluntly, surrendering information about the identity of the atomic electrons in the molecular configuration leads directly to a significant and observable component of the binding energy, thus posing an equivalence between information and energy far more stark than that implicit in the second law of thermodynamics.

Using similar formalism, we may represent the state function of two interacting individuals, or of an interacting individual and physical system, by a composite state function ψ^{ab} whose behavior characteristics also differ significantly from those of the separated systems, i.e.,

$$s_i^{ab} = s_i^a + s_i^b + \Delta s_i^{ab} + \Delta s_i^{ba}$$

where s_i^a denotes the "normal" response of the first individual to the prevailing situation, and s_i^b that of the second individual or of the physical system, and Δs_i^{ab} and Δs_i^{ba} denote modifications of those behaviors arising because the two systems are strongly interacting during the observed situation.

As a specific example, to apply this approach to remote perception experiments we could denote the percipient by ψ^p, the agent by ψ^a, and the experimental protocol by the mathematical operator P. In the absence of interaction between the percipient and agent, each would have certain "normal" reactions to the experimental situation, p_i^p, p_i^a, derived from the eigenvalue relations:

$$P\psi_i^p = p_i^p\psi_i^p$$
$$P\psi_i^a = p_i^a\psi_i^a$$

i.e., the percipient would perceive nothing about the target that was not accessible to his normal perceptual modes, and the agent would react to the target under no influence from the percipient.

However, if the percipient and agent are strongly enough interacting to require a new "molecular" wave function, ψ^{pa}, "paranormal" terms will appear in their response patterns:

$$p_i^{pa} = p_i^p + p_i^a + \Delta p_i^{pa} + \Delta p_i^{ap}$$

where in Δp_i^{pa} we may accommodate the anomalous acquisition of information about the target, and in Δp_i^{ap} the commonly reported experience of the agent of having his attention attracted to specific details he would "normally" have ignored.

Application of this formalism to a PK experiment proceeds in a similar fashion. Here we might represent the experimental operator (person) by ψ^o, and the experimental device by ψ^d. Again the experimental protocol is represented by a mathematical operator, K. In the absence of major interaction, the device behaves "normally":

$$K\psi_i^d = k_i^d\psi_i^d$$

and the operator has the "normal" psychological experiences:

$$K\psi_i^o = k_i^o\psi_i^o.$$

But if the operator and the device are in some state of resonance, ψ^{do}, each behaves somewhat differently, i.e.,

$$k_i^{do} = k_i^d + k_i^o + \Delta k_i^{do} + \Delta k_i^{od}.$$

The anomalous modification in the behavior of the system Δk_i^{do} is termed PK; Δk_i^{od} accomodates any paranormal psychological reactions of the operator.

Development of further illustrations of this general method here would be inappropriate and would require more detailed specification of the nature of the consciousness wave functions, their functional form, their proper "soft" coordinates, and the interpretation of their quantum numbers. Some of this has been attempted, along with various other applications to the psychic domain, and is available in a reference [210]. The point of exposition here is largely a philosophical one: namely, the "paranormal" effects emerge as a consequence of the comparison of the behavior of an interacting system with that of its separated components.

Quantum mechanics may have quite another analytical precedent to contribute to the representation of psychic phenomenology. Clearly, much of the psychic research data will continue to be acquired and processed in statistical form, using established statistical methods. Yet, all statistical models ultimately trace back to certain fundamental probability rules for the elemental systems involved. For the statistical models to be viable, these probability rules must a) exist; b) be known; and c) be analytically tractable. At present, virtually all processing of psychic research data presumes the applicability of classical statistics, yet the basic probability rules for the elemental processes are, in point of fact, unknown [211], [212].

It may fortuitously be the case that much of psychic process can be adequately treated as marginal deviation from classical chance behavior. In some cases, however, it may be necessary to invoke categorically different statistics, tracing back to fundamentally different probability rules, to deal with the effects. The quantum mechanical precedents of relevance, of course, are

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the two systems of quantum statistics, i.e., the Fermi–Dirac and Bose–Einstein systems, which are based upon the phase-space population rules for half-integer and integer spin particles as imposed by the Pauli principle, i.e., the wave function symmetry requirements. For most common physical systems, these quantum statistics conveniently degenerate into the classical form, but for certain special situations, e.g., the specific heats of metals and certain radiation properties, their full application is essential.

By analogy, one could postulate that physical reality, as perceived and influenced by human consciousness, actually plays by more elaborate probability rules than commonly attributed to it, and hence strictly requires a more complex statistical mechanics. Thus, in this view, the processes commonly regarded as "normal" would be those for which the "classical" approximation to this more complex system is adequate; small-scale psychic effects would then comprise those displays of minor deviations of the complex statistics from the "classical" limit; the more drastic phenomena—poltergeists, levitations, metal bending, etc.—would presumably become explicable only in terms of the full statistical formulation.

Holistic Models

Yet more extreme in conceptual difficulty is a body of contention that psychic processes are inseparably holistic, and that no model rooted in any sector of established science can adequately represent them [213]. In particular, the suggestion is offered that psychic processes are manifestations of the interdiffusion of the analytical, scientific world with the creative aesthetic world, and thus to represent them effectively it will be necessary to combine the philosophical perspectives and techniques of both domains. To resort to metaphor for illustration once again, one could ponder such analogies from the common physical world as the interface regions between the sea and land, where the diffuse patterns of the ocean wave structure meet the solid promontories and sloping beaches of the coast to produce the crashing breakers and hissing foam of seashore phenomena, or the overrunning of a warm, moist atmosphere by a cold-front of drier, cooler air to initiate the striking electrical and acoustical phenomena of the summer thunderstorm.

By whatever analogy it may be illustrated, theoretical representation of the interpenetration of causal physical mechanics with creative conscious process must be a formidable undertaking, yet not totally without precedent or allied effort. For examples, interest is just now growing in the humanistic psychology community in the analytical study of human creativity [214]–[216], and, on the other side of the interface, a few physicists are beginning to muse openly about the role of aesthetics in subnuclear and cosmological physical behavior [217]–[220]. In yet a different arena, certain futurists are now examining the interplay of aesthetic and functional human needs and values in the evolution of social and political structure [221]. To be sure, none of these has produced much in the way of analytical formalism, but the peculiarity, magnitude, and significance of the interpenetration effects are being acknowledged.

Any summary assessment of the status of physical models of psychic phenomena should properly begin with reiteration of the opening statement: none of the approaches outlined above has yielded anything approaching a functional theoretical basis. Yet, the ensemble of empirical experimental experience seems to suggest that certain of the conceptual and perceptual characteristics underlying those formalisms may be relevant to ultimate representation of such processes. Specifically, they suggest that the following rather general hypotheses may be worthy of more detailed examination:

1) The phenomena may be inherently statistical, rather than directly causal, and we may be observing them "on the margin." That is, the observed phenomena may represent marginal changes from normal behavior on a very grand scale and with fluctuation times which tax human observational capability. It also may be necessary to deal with more complex statistical mechanics, appropriate to more involved basic probability rules, to represent the most drastic effects.

2) Just as human consciousness has the ability to extract information from an external system, e.g., by observing it, that consciousness may also have the ability to project information into it, e.g., by ordering random processes.

3) Quantum mechanics may be more than a system of physical mechanics; it may be a more fundamental representation of human consciousness and perception processes, and the empirical pillars of this formalism, such as the uncertainty principle, the exclusion principle, the indistinguishability principle, and the wave/particle dualities may be as much laws of consciousness as laws of physics.

4) Psychic processes may be inherently holistic, and thus the ultimate model may need to integrate both the scientific and the aesthetic aspects in order to identify the sources of the phenomena. That is, psychic processes may be manifestations of the intersection of the analytical, scientific world with the creative, aesthetic world, and thus, to represent them effectively, it may be necessary to integrate both perspectives without sacrificing the integrity of either.

Clearly, any of these intuitions will have to be developed in far more philosophical and analytical detail before a trenchant theoretical model can emerge, but at this primitive stage it is probably stimulating to consider a few such radical possibilities, along with more prosaic explications. Changes of this magnitude in representation of human perceptual reality inevitably, and properly, would be attended by much philosophical recalcitrance and agony, but the broader personal and collective insights that could derive from legitimate efforts to bridge the analytical/aesthetic interface could be of at least corresponding benefit.

The Negative Side

Contemporary criticism of psychic research and rejection of the phenomena it purports to demonstrate tend to focus on a number of specific objections, each of which has some degree of validity and merits some thought in any balanced assessment of the topic [118], [178], [222]–[224]. The most commonly cited concerns include:

1) demonstrable fraud;
2) naivete of technique, including inadequate controls, faulty equipment, sensory cuing of participants, other experimenter biases, selective treatment of data, improper statistical methods, and general experimental and theoretical incompetence;
3) little improvement in comprehension over many years of study;
4) absence of adequate theoretical models;
5) suppression of negative results;
6) poor experimental replicability;
7) elusiveness of effects under close scrutiny;

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

8) sensitivity of results to participants, attitudes, and laboratory ambience;
9) tendency for many results to be only marginally significant compared to chance expectation;
10) inconsistency with prevailing "scientific world view;"
11) contradictory to personal psychology, philosophy, theology, or "common sense."

Obviously this list runs a gamut from rather technical and procedural objections, through phenomenological inconsistencies, to rather categorical and subjective rejections, and only a few of these issues can be constructively addressed here.

Unfortunately, but undeniably, one or the other of the first two judgments may legitimately be applied to a large body of the propounded results. By its nature, the field is immensely vulnerable to fraudulent exploitation and naive gullibility, and such have indeed occurred to a distressing degree. It is also true that the topic has attracted a disproportionate share of less than fully competent researchers, and that it presents extraordinary pitfalls for even the most disciplined scholars. Yet despite the substantial validity of these claims, and the suspicion they inevitably cast on all other results, it does not seem that they should predicate categorical rejection of the entire field. Rather, the vulnerable cases should be patiently ferreted out using obvious scholarly criteria, and only those efforts surviving such scrutiny used for scientific insight and judgment.

The lack of definitive progress toward comprehension of the phenomena and the absence of viable theoretical models have already been acknowledged in the foregoing text, and although these assessments should perhaps be qualified by the relatively minute integrated investment of resources made in this field in comparison to many of the more favored areas of science, they nonetheless constitute legitimate concerns about the ultimate tractability of the field. It should perhaps also be noted that, despite the prolonged effort, it is only very recently that more sensitive and powerful experimental equipment and data processing techniques have been brought to bear, and equally recently that more sophisticated physical formalisms have been invoked, and that these have had a much briefer opportunity to render the phenomena into comprehensible terms.

The subsequent five objections, 5)-9), are more specific and substantive, and merit examination from two orthogonal points of view. Namely, to what extent do such characteristics indeed invalidate the results, and conversely, to what extent might they illuminate the basic nature of the phenomena? With respect to suppression of unfavorable results, there has undoubtedly been some tendency in this field, as in most others, to advance positive or definitive findings more enthusiastically than negative or equivocal ones. Indeed, this paper has been guilty of the same bias. In an effort to provide concise representation of the style and substance of psychic research and of the nature of the effects it can produce, we have tended to invoke as illustrations some of the more successful and familiar pieces of work without balancing the presentation with comparable examples of the negative or equivocal results that are regularly acquired in these efforts.

To the credit of the psychic research community, it has officially encouraged thorough and objective reporting of negative data, and much of these indeed appear regularly in its established journals [225]-[228], with a number of consequent benefits. First, beyond adding credibility to the body of positive results, such data compound to provide some quantitative index of the ratio of positive to negative yields in a given class of experiment, thereby contributing to a broader sense of the grand statistics of the phenomena. In addition, documentation of the specific conditions prevailing in unsuccessful experiments may be helpful in excluding irrelevant parameters from further consideration, and in identifying and reducing counterproductive influences. Perhaps most pointedly, however, the body of negative and equivocal data emphasize that psychic phenomena, if real, are highly irregular and sensitive to intangible influences well beyond current scientific control, and, if their study is to be pursued, this caveat must be accepted *ab initio*, at least for the present.

A similar interpretation also applies to the irreplicability complaint, and to the three following it. Without question, the dominant experimental frustration in this field is the inability to replicate on demand previously observed paranormal effects, not only at other laboratories, but even in the original facility, using the original participants, under apparently identical experimental circumstances [172], [173], [229]-[231]. This ubiquitous characteristic has precipitated major philosophical excursions which are well beyond our capacity to review here. Only briefly, four possible categorical interpretations have been advocated:

a) The phenomena are illusory.
b) The phenomena are rare and bizarre chance occurrences, beyond any hope of regularization.
c) The phenomena are precipitated, at least in part, by psychological and/or physiological factors which are presently beyond experimental control, but which if fully comprehended would conform to established scientific paradigm.
d) The phenomena are inherently statistical, and possibly quantum mechanical, on a macroscopic scale, thus manifesting themselves with finite but fractional probability on any given occasion.

The latter pair of options are not necessarily mutually exclusive, particularly if one takes a rather generous doubly statistical point of view, namely that the human population embodies a range of capability for engendering such effects and that beyond that, any individual may display a variable range of personal capacity, depending on a variety of environmental, physiological, and psychological factors prevailing at the time.

The evasiveness of the phenomena under carefully controlled and observed study may be the most damning criticism of all, or it may also constitute a valid and illuminating phenomenological characteristic. The tendency of a given preliminary or anecdotal effect to disappear or diminish when the experiment is tightened, or when it is displayed to a skeptical jury of observers, obviously casts major doubt on the scientific integrity of the process. Yet it also brings to mind at least two other processes which, while superficially dissimilar, may not be totally irrelevant, namely artistic creativity, and quantum mechanical measurement as limited by the uncertainty principle. In the former, there should be little quarrel that the creative processes of artistic, musical, or literary composition, or of lofty philosophical thought in general, are not usually facilitated by rigid constraints or by the presence of a body of unsympathetic observers. The importance of favorable ambience and mood for such efforts is intuitively and demonstrably clear, and little creative achievement is likely to occur in overly sterile or hostile environments, a truth Richard Wagner vividly conveyed to his own critics by his portrayal of the fate of young Walther at-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

tempting his avant-garde Trial Song before the assembled Meistersingers. Virtually every creative artist preserves some form of retreat or sanctuary, and even the most rigorous of scientists will concede the role of unstructured mental imagery in enhancing their technical insights.

The analogy of the quantum mechanical measurement process is somewhat more strained, in that it requires generalization of the concept to the macroscopic level of information or energy exchange between two persons or between a person and a physical system. The point will not be developed here, other than to note that if there is any validity to the application of quantum mechanical logic to this class and scale of intellectual/intuitive process, as discussed earlier and in the references, some form of "uncertainty principle" could predicate a limit to the precision with which psychic effects could be observed. More specifically, if the "hard" and "soft" coordinates of representation are canonically conjugate, some form of $\Delta q \Delta p \sim h$ rule may apply, so that attempts at excessive precision in specification of a psychic effect could dissipate its cause, and vice-versa [210].

The final two reservations regarding inconsistency of the phenomena with established scientific and personal views, while constituting powerful professional and personal discriminators and properly predicating great caution and discipline in venturing into any anomalous field, also cannot be allowed total veto authority if new domains of conceptual experience are ever to be challenged. In responding to a critic of an earlier paper who stood on these points I wrote, perhaps too floridly,

> Authoritarianism such as this encourages established knowledge to sit smugly on its duff and categorically reject all new evidence that does not support or fill in its contemporary "world-view compatability criterion"–whatever that is. Worst of all, it stifles the most precious attribute of human consciousness, the yearning for ever new, ever higher wisdom that has driven the mind and spirit of man to evolve upward, rather than merely to replicate [232].

More persuasive to this issue, however, would be a simple historical count of the number of leaps of scientific insight, from Aristotle to this day, which would have been, and in most cases were, for a time, rejected on the basis of these criteria. Curiously, it has often been those giants of science who with soaring insight and courage of conviction violated such tenets to lead their fields to new plateaus of understanding, who also, in a later day, led the recalcitrance of the establishment against comparably sacrilegious visions of their successors, while still endorsing in general terms the importance of visionary thought. Galileo, early champion of scientific methodology and revolutionary concepts in terrestrial and celestial mechanics against vicious dogmatic opposition, rejected Kepler's elliptical orbits as "occult fantasy"; Thomas Young, whose brilliant interference experiments finally established the wave character of light, contended with Fresnel's theoretical formulations of the same processes; Ernst Mach disputed relativity and atomic theory; Rutherford, who showed the world the nuclear atom, dismissed any practical significance for nuclear energy; Lavoisier and Ostwald disputed atomic theories of chemistry; D'Alembert opposed probability theory; Edison discounted alternating current; Lindberg despaired of Goddard's rocketry; and Albert Einstein retained an enduring uneasiness about quantum theory despite his many contributions to its evolution [233]–[235]. Incidentally, the same Albert Einstein who would invoke the establishment criteria against Upton Sinclair's clairvoyance claim:

> ...the results of the telepathic experiments carefully and plainly set forth in this book stand surely far beyond those which a nature investigator holds to be thinkable [152].

could in quite another tone testify eloquently to the importance of the aesthetic dimension in creative science:

> The most beautiful and most profound emotion we can experience is the sensation of the mystical. It is the sower of all true science. He to whom this emotion is a stranger, who can no longer wonder and stand rapt in awe, is as good as dead. To know that what is impenetrable to us really exists, manifesting itself as the highest wisdom and the most radiant beauty which our dull faculties can comprehend only in their most primitive forms—this knowledge, this feeling is at the center of true religiousness.
>
> The cosmic religious experience is the strongest and noblest mainspring of scientific research [236].

Individually and collectively, these critical challenges to psychic research raise valid concerns which merit deliberate attention and predicate great caution, and can also help to illuminate some of the subtle phenomenological features. However, they can perform these functions well only if they themselves are informed, reasoned, and fair. Regrettably, from my reading of the critical literature, this has not invariably been the case, and instances of naiveté, selective representation of data and protocol, and excessive generalization also appear therein.

The role of the critic in psychic research is a most essential one, perhaps more so than in any field of scholarship yet broached. When the criticism is based in fact and experience and is objective and fair, it can instill healthy discipline in the study of this or any other difficult field and ensure that the fundamental requisites of scientific methodology, e.g., dispassionate rigor, humility in the face of observations, limitation on extrapolation of results, and openness of mind will prevail in the search. But it is equally essential that the process of criticism play by these same rules. If it violates any of them, if it lapses into categorical rejection, guilt by association, or sloppy logic, it can become as suspect as the object of its complaint, and thus fail in its proper role [237].

Implications

Despite their compounded length, the foregoing historical outline, survey of contemporary activity, and selected examples of ongoing research and theoretical efforts are still far from adequate to convey the full essence of this complex and contradictory field. Fortunately an extensive body of reference literature exists, including a number of comprehensive general volumes, whereby an interested reader may flesh out this sketch and extend it to many aspects not broached here [6], [44], [45], [56], [57], [92], [130], [238]–[242]. Hopefully, such more thorough study would tend to confirm the general impressions of the status and prospects of the field conveyed above. To restate these in summary, it appears that once the illegitimate research and invalid criticism have been set aside, the remaining accumulated evidence of psychic phenomena comprises an array of experimental observations, obtained under reasonable protocols in a variety of scholarly disciplines, which compound to a philosophical dilemma. On the one hand, effects inexplicable in terms of established scientific theory, yet having numerous common characteristics, are frequently and widely observed; on the other hand, these effects have so far proven qualitatively and quantitatively irreplicable, in the strict scientific sense, and appear to be sensitive to a variety of psychological and environmental factors that are difficult to specify,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

let alone control. Under these circumstances, critical experimentation has been tedious and frustrating at best, and theoretical modeling still searches for vocabulary and concepts, well short of any useful formalisms.

Given these difficulties, what then are the motivations, if any, to proceed? As for most speculative topics, three potential generic benefits could be considered:

a) acquisition of fundamental knowledge,
b) practical applications,
c) humanistic benefits.

In this particular field, basic knowledge might accrue in two ways—the attainment of new scientific information in the usual sense, and the broadening of scientific methodology to deal more effectively with irregular phenomena of this type. In other words, study of this topic not only might provide certain phenomenological answers, but also might serve to broaden the context in which science can formulate its questions. New mechanisms for transfer of information or energy might be identified, or broader understanding of those properties, and how they are perceived and measured might emerge. The latter half of the opportunity clearly is a major challenge to science, but hardly a new one. William James posed it rather bluntly some eighty-five years ago:

> The spirit and principles of science are mere affairs of method; there is nothing in them that need hinder science from dealing successfully with a world in which personal forces are the starting point of new effects. The only form of thing that we concretely have is our own personal life. The only completed category of our thinking, our professors of philosophy tell us, is the category of personality, every other category being one of the abstract elements of that. And this systematic denial on science's part of personality as a condition of events, this rigorous belief that in its own essential and innermost nature our world is a strictly impersonal world, may, conceivably, as the whirligig of time goes round, prove to be the very defect that our descendants will be most surprised at in our boasted science, the omission that to their eyes will most tend to make it look perspectiveless and short [24].

The potential applications of psychic process are best considered with conservatism and restraint, especially given the tendency of certain elements of the public media and private exploiters to extrapolate the possibilities far beyond any demonstrated accomplishments. Clearly, the process of remote perception described earlier, along with other forms of clairvoyance, could hold some potential interest for intelligence agencies, law enforcement units, and any other activity relying on surveillance, as well as for archaeological searches, natural resource prospecting, and the like, and such operations have indeed engaged in empirical efforts to evaluate the efficacy of such strategies in their particular domains. From a strictly engineering standpoint, however, the potential efficiency and precision of such tactics are unclear, given their apparent tendency to trade more effectively in impressionistic generalities than in analytical detail.

Low-level PK effects, such as the REG distortions indicated above, could have more pervasive implications for high technology. If, for example, the basic functions of microelectronic elements could be even slightly disturbed by intentional or inadvertent intervention of human consciousness, it would seem important to obtain some assessment of the potential magnitude of such effects, and of the factors favoring or inhibiting such interference, before much more elaborate integrated circuit arrays, graphic display systems, and other sensitive man/machine interfaces are committed to delicate or critical operations. To focus our assessment of such possibilities, we are now examining PK disturbance of the memory function of a single microelectronic chip [243]. If the indications that psychological and environmental factors bear on the precipitation of such effects are sustained, it may be necessary to expand consideration of such parameters beyond the usual scope of human factors engineering, especially in situations involving high psychological stress.

The potential humanistic benefits of better comprehension of psychic phenomena could be addressed on either a personal or social level, but to do so in any detail would far exceed the purview of this article. Again, extensive references on various facets of the issue abound [244]-[252]. Ultimately, most of these philosophical excursions arrive at the same monumental question, namely whether convincing demonstration of the capability of human consciouness to influence its reality to a significant degree would substantially alter individual and collective perception of the human state, its value system, and its behavior pattern, and thereby facilitate its evolution to a higher life form. Such projections have been offered from a variety of perspectives. Engineer/futurist Willis Harman forecasts an "inner experience" paradigm:

> Just as conventional science depends upon a prior consensus on how knowledge of the sense-perceived world shall be publicly tested and validated, so the complementary paradigm will have to include consensus on how knowledge relating to the world of inner experience shall be publicly tested and validated.
>
> Its essential characteristic would be that consciouness and its contents are primary data, rather than being secondary and derivative as in the conventional paradigm. Where the conventional paradigm involves reductionistic models the complementary paradigm would add holistic models; where the first employs deterministic (or stochastic) explanations of events the second would add teleological, purpose-recognizing explanations; where the first is little involved with matters of values and meaning, the second finds these of central concern; where the first is dominated by technology-focused values of prediction and control, the second would tend to value understanding relating to human well-being, development and evolution [221].

Biologist/immunologist Jonas Salk phrases it more in terms of a resonance of human volition with natural processes:

> Man has come to the threshold of a state of consciousness, regarding his nature and his relationship to the Cosmos, in terms that reflect 'reality.' By using the processes of Nature as metaphor, to describe the forces by which it operates upon and within Man, we come as close to describing 'reality' as we can within the limits of our comprehension. Men will be very uneven in their capacity for such understanding, which, naturally, differs for different ages and cultures, and develops and changes in the course of time. For these reasons it will always be necessary to use metaphor and myth to provide 'comprehensible' guides to living. In this way, Man's imagination and intellect play vital roles in his survival and evolution [249].

And philosopher/paleontologist Teilhard de Chardin states his hope in terms of a collective consciousness of the human race:

> Thus we find ourselves in the presence, in actual possession, of the super-organism we have been seeking, of whose existence we were intuitively aware. The collective mankind which the sociologists needed for the furtherance of their speculations and formulations now appears scientifically defined, manifesting itself in its proper time and place, like an object entirely new and yet awaited in the sky of life. It remains for us to observe the world by the light it sheds, which throws into astonishing relief the great ensemble of everyday phenomena with which we have always lived, without perceiving their reality, their immediacy or their vastness [253].

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

On a somewhat less lofty, but possibly more functional level, a recent lengthy study prepared for the House of Representatives Science and Technology Committee, in a section encouraging serious assessment of further research on "the physics of consciousness," stated that recent experiments:

> suggest that there exists an 'interconnectiveness' of the human mind with other minds and with matter; . . . that the human mind may be able to obtain the information independent of geography and time

and later concluded that:

> . . . a general recognition of the degree of interconnectiveness of minds could have far-reaching social and political implications for this nation and the world [254].

The details and tones of these visions clearly are matters of individual heritage, experience, and intuition, but the messages share a common theme: the next stage in human evolution may involve expansion and interconnection of human consciousness, features clearly central to the psychic concept.

Beyond the difficulty of the phenomena and potential benefits of the knowledge, a more prosaic factor to be considered in contemplating further psychic research is the requisite cost of the effort. To date, such experimentation has been extraordinarily inexpensive by usual scientific standards. The primitive level of comprehension, the lack of organized interest within the established scientific communities, and the Bohemian status of many of the investigators have predicated projects of very low budgets compounding to a total annual national investment of the order of one million dollars or less [255]. Clearly, if more incisive progress is to be attempted, some increase in the sophistication and interaction of the principal programs must be funded, but it is most unlikely, and for the time probably undesirable, that this will consume any magnitude of resources comparable with, or distractive from, the better established research domains. A comparable statement could be made with respect to commitments of the requisite scholarly personnel.

If on the basis of such cost/risk/benefit considerations a modest ongoing program of research seems justified, it remains to consider the selection of topics, experimental styles, and evaluation criteria which would optimize the effort at this stage. In addition to the obvious desiderata that the specific experiments be clearly posed, conceptually simple, lend themselves logistically to rigorous, tightly controlled protocols, and focus on the more tractable and potentially applicable effects, three more specific recommendations could be offered.

First, given the irregularity of the phenomena, their possible dependence upon a broad range of physical, psychological, physiological, and environmental parameters, and their tendency to display effects as marginal deviations from some "normal" distributions, some premium should probably be placed on the capacity of the experiments for large data base accumulation and processing. While less prolific studies may continue to provide interesting anecdotal effects and suggest procedures for more detailed programs, it is unlikely that much correlation of those effects with pertinent prevailing parameters can be achieved without large quantities of data. More specifically, the favored experiments should 1) deal with processes found to have relatively high intrinsic yields; 2) employ equipment and protocols which permit data acquisition at rapid rates; and 3) have access to computational equipment and software which allow storage of large data arrays and processing of that data in many selective cross-concatenations. As an example, the latest refinement of the REG experiment described earlier allows acquisition, storage, and primary processing of several hundred experimental trials ($\sim 10^5 - 10^6$ bits) per hour, and subsequent concatenation of all previously stored data by any permutation of ten parametric indices, e.g., sampling rate, manual/automatic, volitional/instructed, operator characteristics, etc. For any such parametric explorations to be indicative, the pertinent data subset must be adequately large to display any systematic deviation from the baseline distribution beyond the statistical noise background. This we have found to require at least several thousand trials; the implication for the total data base is clear.

This very large data capability inevitably predicates a sophistication in equipment and software which, beyond the initial and operating expenses, may introduce some undesirable effects of complicating the phenomenological processes and clouding the experimental ambiences. Specifically, one needs to consider whether any observed effects, in either the PK or ESP categories, still trace unequivocally to the primary physical processes, or whether there now might arise confounding interactions with elements of the data collection and processing equipment or techniques. Related to this concern are possible uncertainties of the participants in defining and focusing on the primary tasks.

Reservations of this sort lead directly to a second general recommendation for effective psychic experimentation. Namely, if the phenomena derive to any significant degree from conscious or subconscious processes of the human mind, it is important that such not be inhibited or excessively complicated by the design and operation of the experiments. For this and numerous other reasons, it is probably essential in planning and implementing the experimental programs to include the insights, interpretations, and intuitions of the human operators, especially those who have demonstrated some success in the generation of the phenomena. It is quite possible that the difference between a sterile experiment and an effective one of equal rigor lies as much in the impressionistic aspects of its ambience and feedback as in the elegance of its instrumentation, and the former need to be well-tuned to the participants who are asked to function as components of the experimental system. On occasion, there seems to have been some tendency in this field to treat the experimental participants in rather perfunctory fashion, discounting any insights they might offer on the tasks at hand. If one subscribes at all to the concept of the phenomena emerging from some interpenetration of analytic and intuitive processes as suggested by the holistic models, there would seem no better place to combine perceptions and insights from these two domains than in the design, operation, and interpretation of the experiments addressed to illumination of the interface.

Finally, it seems most evident that given the intrinsic transdisciplinary nature of the business, research on this topic in any established sector should become much more communicative and interactive across traditional scholarly boundaries if it is to have any hope of rendering the phenomena into comprehensible and serviceable terms. This cross-talk cannot be limited to naturally contiguous fields, à la the usual exchange between physicist and engineer, or between psychologist and sociologist. As its lengthy heritage illustrates, in this domain the interests and insights of the theologian, philosopher, statistician, technologist, hard scientist, and creative holistic thinker are all potentially valid, and need to be melded in scholarly symbiosis and common respect. No insular approach is likely to prevail.

This requisite has implications for the staffing of particular projects, for the institutional environment in which they operate, and for the professional societies and publications which choose to attend to this topic. Individual laboratory personnel

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

groups should comprise a broader range of experience and insight than the conventional hierarchy of technical specialists, and their cognizance of other contemporary work should be broader. The institution housing that laboratory needs to display considerable tolerance and support for the unusual tone and special requirements of the research and not force its conformity to established scientific subdivisions and research styles. Likewise, the professional community at large cannot at this time profitably ask for total adherence to its own reductionistic superstructure, but can only inquire dispassionately regarding the respective implications of this conglomerate field for the traditional areas of endeavor.

In this last regard, I should like to express my personal respect for this particular Society, and for this particular Journal, for the openness and generosity of spirit with which they have solicited and presented the results of legitimate scholarly effort in this difficult field. Their attitude could well stand as a model for other institutions and organs in dealing with this topic or with any other present or future projective area of human inquiry.

Acknowledgment

The author would like to express his indebtedness to his research colleague, Ms. B. J. Dunne, who has provided immense logistical help in the preparation of this manuscript and who, as laboratory manager of the Princeton Engineering Anomalies Research program, has played a primary role in the generation and interpretation of much of the data reported herein. Dr. R. D. Nelson also contributed heavily to the experimental program and to the refinement of the manuscript. Prof. W. H. Surber designed much of the REG circuitry and software, and wrote the original programs for the remote perception analytical judging procedure.

The program has also benefitted immensely from the interests and efforts of several present and past students and staff, most notably C. K. Curry (Dunham), E. G. Jahn, T. A. Curtis, and I. A. Cook.

References

[1] M. Ebon, "A history of parapsychology," in *Psychic Exploration*, E. D. Mitchell *et al.*, J. White, Ed. New York: Putnam, 1974.
[2] A Gould, *The Founders of Psychical Research*. New York: Schocken, 1968.
[3] J. B. Rhine and associates, *Parapsychology from Duke to FRNM*. Durham, NC: Parapsychology Press, 1965.
[4] H. Carrington, *The Story of Psychic Science (Psychical Research)*. New York: Ives Washburn, 1931.
[5] R. C. LeClair, Ed., *The Letters of William James and Théodore Flournoy*. (Foreword by Gardner Murphy.) Madison, WI: University of Wisconsin Press, 1966.
[6] B. B. Wolman, Ed., *Handbook of Parapsychology*. New York: Van Nostrand Reinhold, 1977.
[7] D. S. Rogo, *Parapsychology: A Century of Inquiry*. New York: Taplinger, 1975.
[8] F. Hartmann, *Paracelsus: Life and Prophecies*. Blauvelt, NY: Rudolf Steiner Publications, 1973, pp. 103-131.
[9] M. Bell, "Francis Bacon: Pioneer in parapsychology," *Int. J. Parapsychol.*, VI, no. 2, pp. 199-208, Spring, 1964.
[10] J. Glanvill, *Saducismus Triumphatus*, Introduction by C. O. Parsons, Facsimiles edit. Gainesville, FL: Scholars' Facsimiles and Reprints, 1966. (Originally published in 1689.)
[11] R. Haynes, *Philosopher King–The Humanist Pope Benedict XIV*. London, England: Weidenfeld & Nicholson, 1970.
[12] C. Richet, "La suggestion mentale et le calcul des probabilités," *Rev. Philosoph.*, 18, pp. 608-674, 1884.
[13] C. Richet, "Further experiments in hypnotic lucidity or clairvoyance," in *Proc. Soc. for Psychical Research*, 6, pp. 66-83, 1889.
[14] E. Swedenborg, *Divine Love and Wisdom*. New York: Swedenborg Foundation, Inc., 1979.
[15] G. Trobridge, *Swedenborg, Life and Teaching*. New York: Swedenborg Foundation, Inc., 1976.
[16] I. Ross, *The President's Wife, Mary Todd Lincoln*. New York: Putnam, 1973.
[17] F.W.H. Myers, *Human Personality and the Survival of Bodily Death*, (Abridged and edited by S. Smith.) Secaucus, NJ: University Books, 1961. (Originally published in 1903, 2 vols.)
[18] W. F. Barrett, "Appendix to the report on thought-reading," in *Proc. Soc. for Psychical Research*, 1, pp. 47-64, 1882.
[19] O. J. Lodge, "An account of some experiments in thought-transference," in *Proc. Soc. for Psychical Research*, 2, pp. 189-200, 1884.
[20] The American Society for Psychical Research, Inc., 5 W. 73 St., New York, NY 10023.
[21] W. James, *Human Immortality: Two Supposed Objections to the Doctrine*. Boston, MA: Houghton Mifflin, 1898.
[22] ——, "Report of the Committee on Mediumistic Phenomena," in *Proc. American Soc. for Psychical Research*, 1, pp. 102-106, 1896.
[23] G. Murphy and R. O. Ballou, *William James on Psychical Research*. New York: Viking Press, 1969.
[24] W. James, "Psychical research," in *The Will to Believe and Other Essays in Popular Philosophy*. New York: Longmans, Green, and Co., 1897.
[25] ——, "Review of human personality and its survival of bodily death by F.W.H. Myers," in *Proc. Soc. for Psychical Research*, 18, p. 23, 1903.
[26] S. Freud, "A note on the unconscious in psycho-analysis," in *Proc. Soc. for Psychical Research*, 26, pp. 312-318, 1912.
[27] ——, *Studies in Parapsychology*. New York: Collier Books, 1963.
[28] C. G. Jung, *Synchronicity*. Princeton, NJ: Princeton University Press (Bollingen Series #20), 1973.
[29] ——, *On the Psychology and Pathology of So-Called Occult Phenomena*. Princeton, NJ: Princeton University Press (Bollingen Series # 20), 1957. (Originally published in 1902, doctoral dissertation, University of Basel.)
[30] ——, *Memories, Dreams, Reflections*. (Recorded and edited by A. Jaffe, translated by Richard and Clara Winston.) New York: Vintage Books, 1965, p. 137.
[31] W. Pauli and C. G. Jung, Eds., *The Interpretation of Nature and the Psyche*. Princeton, NJ: Princeton University Press (Bollingen Series #51), 1955.
[32] J. Ehrenwald, "Psi, psychotherapy, and psychoanalysis," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 529-540.
[33] M. Ullman, "Psychopathology and psi phenomena," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 557-574.
[34] J. G. Pratt, J. B. Rhine, C. E. Stuart, and B. M. Smith, *Extra-Sensory Perception After Sixty Years*. New York: Holt, 1940.
[35] J. B. Rhine, *New World of the Mind*. New York: William Morrow and Co., 1953.
[36] J. B. Rhine, Ed., *Progress in Parapsychology*. Durham, NC: Parapsychology Press, 1971.
[37] L. E. Rhine, *Mind Over Matter: Psychokinesis*. New York: Macmillan, 1970.
[38] J. B. Rhine, "Psi and psychology: Conflict and solution," *J. Parapsychol.*, 32, pp. 101-128, 1968.
[39] *Eur. J. Parapsychol., Int. J. Parapsychol.* (defunct), *J. Parapsychol., J. Amer. Soc. Psychical Research, J. Soc. for Psychical Research, Parapsychol. Rev., Proc. Soc. for Psychical Research.*
[40] US Universities/Colleges: Chicago; Colorado; Columbia; CUNY; Delaware; Denver; Drexel; Duke; Finch; Harvard; Illinois; John F. Kennedy; Kent State; Michigan; Missouri; Mundelein; Newark College of Engineering; New School for Social Research; NYU; North Carolina; Oakland; Pittsburgh; Princeton; St. John's; St. Joseph's; South West Minnesota; Syracuse; UC/Berkeley; UC/Davis; UC/Irvine; UC/LA; UC/Santa Barbara; Virginia; Wesleyan; West Florida; West Georgia; Wisconsin; Yale. Foreign: Amsterdam; Andhra; Argentina; Benares; Bonn; Cambridge; Chile; Copenhagen; Edinburgh; Freiburg; Ghana; Halifax; Iceland; Jung Inst.; Leningrad; Leyden; London; Loyola; Lund; McGill; Munich; Mysore; New South Wales; Oxford; Paris; Peking; Rhodes; South Africa; Tasmania; Tel Aviv; Utkal; Utrecht; Warsaw; Waterloo; Zurich.
[41] List available upon request.
[42] Research Institutions: Amer. Soc. for Psychical Research, New York; Ballistic Research Labs., Aberdeen, MD; Foundation for Research on the Nature of Man, Durham, NC; Inst. for Noetic Sciences, San Francisco, CA; Menninger Foundation, Topeka, KS; Midwest Psi Research Inst., Chicago, IL; Mind Science Foundation, San Antonio, TX; Parapsychology Foundation, NY; Psychical Research Foundation, Durham, NC; Psychophysical Research Labs, Princeton, NJ; Science Unlimited Research Foundation, San Antonio, TX; SRI International, Menlo Park,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

CA; Society for Psychical Research, London; Forschungsinstit für Psychotronik, West Berlin; Matsui Hospital, Kasahara, Japan; Inst. of Pharmacology, Polish Academy of Sciences, Krakow; Shanghai Municipal Human Exception Functions Research Society. Corporations: Airesearch Company of California; Boeing Scientific Research Laboratories.

[43] J. G. Pratt, "Soviet research in parapsychology," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 883-903.

[44] E. K. Naumov and L. V. Vilenskaya, *Bibliographies on Parapsychology (Psychoenergetics) and Related Subjects.* Arlington, VA: Joint Publications Research Service, 1972. (Originally published in Moscow, 1971.)

[45] E. K. Naumov and L. V. Vilenskaya, *Bibliography of Parapsychology (Psychotronics, Psychoenergetics, Psychobiophysics) and Related Problems.* (Literature published in the USSR), 1979 (in manuscript).

[46] E. C. Wortz, A. J. Bauer, R. F. Blackwelder, J. W. Eerkins, and A. J. Sour, "An investigation of Soviet psychical research," in *Proc. Special Session: The State of the Art in Psychic Research; Electro 77 Professional Program.* New York: Apr. 19-21, 1977.

[47] A. J. Lewis, *A Report.* Document EW-76-011, Airesearch Manufacturing Co. of California, January 1976. (A supplement to Wortz *et al.*, #46 above.)

[48] M. Johnson (Chairman) *et al.*, "Symposium: Current directions in European parapsychology," in *Research in Parapsychology 1978.* Metuchen, NJ: Scarecrow Press, 1979, pp. 1-10.

[49] S. Ostrander and L. Schroeder, *Psychic Discoveries Behind the Iron Curtain.* Englewood Cliffs, NJ: Prentice-Hall, 1970.

[50] L. L. Vasiliev, *Experiments in Distant Influence* (Edited with Introduction by A. Gregory). New York: Dutton, 1976.

[51] T. Moss, "Psychic research in the Soviet Union," in *Psychic Exploration*, E. D. Mitchell *et al.*, J. White, Ed. New York: Putnam, 1974, pp. 469-486.

[52] R. A. McConnell, "Parapsychology in the USSR," *J. Parapsychol.*, 39, pp. 129-134, 1975.

[53] V. P. Zinchenko, A. N. Leont'yev, B. F. Lomov, and A. R. Luria, "Parapsychology: Fiction or reality?" *Voprosy Filosofii* (Questions of Philosophy), 9, pp. 128-136, 1973.

[54] Zhang Feng, "New advances made in study of exceptional human functions," Shanghai *Ziran Zazki* (Nature Journal) in Chinese, no. 8, p. 606, Aug. 1980.

[55] J. Palmer, "Attitudes and personality traits in experimental ESP research," in *Handbook of Paraspsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 175-201.

[56] S. Krippner, Ed., *Advances in Parapsychological Research: II, Extrasensory Perception.* New York: Plenum Press, 1978.

[57] K. R. Rao, *Experimental Parapsychology: A Review and Interpretation.* Springfield, IL: Charles C. Thomas, 1966.

[58] G. Schmeidler, "The psychic personality," in *Psychic Exploration*, C. Mitchell *et al.*, J. White, Ed. New York: Putnam, 1974.

[59] H. J. Eysenck, "Personality and extrasensory perception," *J. Soc. Psychical Research*, 44, pp. 55-71, 1967.

[60] G. R. Schmeidler and R. A. McConnell, *ESP and Personality Patterns.* Westport, CT: Greenwood Press, 1973 (originally published 1958).

[61] J. Palmer, "Scoring in ESP tests as a function of belief in ESP. Part I. The sheep-goat effect," *J. Amer. Soc. Psychical Research*, 65, pp. 373-408, 1971.

[62] A Parker, *States of Mind: ESP and Altered States of Consciousness.* New York: Taplinger, 1975.

[63] R. L. Van de Castle, "Sleep and dreams," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 473-499.

[64] M. Ullman, S. Krippner, and A. Vaughan, *Dream Telepathy.* New York: Macmillan, 1973.

[65] I. Strauch, "Dreams and psi in the laboratory," in *Psi Favorable States of Consciousness*, R. Cavanna, Ed. New York: Parapsychology Foundation, 1970, pp. 46-54.

[66] C. Honorton, "Psi and internal attention states," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 435-472.

[67] W. Braud, "Psi conducive conditions: Explorations and interpretations," in *Psi and States of Awareness: Proc. Int. Conf. held in Paris, France, Aug. 24-26, 1977*, B. Shapin and L. Coly, Eds. New York: Parapsychology Foundation, 1978.

[68] E. Dingwall, Ed., *Abnormal Hypnotic Phenomena* (4 Vols.). New York: Barnes and Noble, 1968.

[69] C. Honorton and S. Krippner, "Hypnosis and ESP performance: A review of the experimental literature," in *Surveys in Parapsychology*, R. White, Ed. Metuchen, NJ: Scarecrow Press, 1976, pp. 227-270.

[70] C. Sargent, "Hypnosis as a psi conducive state: A controlled replication study," *J. Parapsychol.*, 4, pp. 257-275, 1978.

[71] C. Tart, "Drug induced states of consciousness," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 500-525.

[72] J. Eccles, "The human person in its two-way relationship to the brain," Invited address to the Parapsychological Association Convention, State University of Utrecht, The Netherlands, Aug. 20, 1976. *Research in Parapsychology 1976.* Metuchen, NJ: Scarecrow Press, 1977, pp. 251-262.

[73] J. Beloff, "ESP: The search for a physiological index," *J. Amer. Soc. Psychical Research*, 47, Sept. 1974.

[74] J. Ehrenwald, "Psi phenomena and brain research," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 716-729.

[75] D. Dean, "The plethysmograph as an indicator of ESP," *J. Soc. Psychical Research*, 41, pp. 351-353, 1962.

[76] J. C. Carpenter, "Intrasubject and subject-agent effects in ESP experiments," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 202-272.

[77] R. A. White, "The influence of experimenter motivation, attitudes, and methods of handling subjects on psi test results," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 273-301.

[78] C. Honorton, M. Ramsey, and C. Cabibbo, "Experimenter effects in extrasensory perception," *J. Amer. Soc. Psychical Research*, 69, pp. 135-149, 1975.

[79] G. Murphy, "Psychical research and human personality," *Proc. Soc. Psychical Research*, 48, pp. 1-15, 1949.

[80] J. B. Rhine, "Conditions favoring success in psi tests," *J. Parapsychol.*, 12, pp. 58-75, 1948.

[81] G. R. Schmeidler and M. Maher, "The non-verbal communications of psi-conducive and psi-inhibitory experimenters," *J. Amer. Soc. Psychical Research*, 75, pp. 241-257, 1981.

[82] W. Crookes, *Crookes and the Spirit World*, collected by R. G. Medhurst. New York: Taplinger, 1972.

[83] B.J.T. Dobbs, *The Foundations of Newton's Alchemy.* Cambridge, MA: Cambridge University Press, 1975, pp. 7-8.

[84] C. Honorton, "Has science developed the competence to confront claims of the paranormal?" in *Research in Parapsychology, 1975.* Metuchen, NJ: Scarecrow Press, 1976, pp. 199-223.

[85] B. Musicante, *Experiments in Psychokinesis: Critique and Proposed Study*, Senior Thesis submitted to the Department of Psychology, Princeton University, May 1979.

[86] R. Jungerman and J. Jungerman, "Computer controlled random number generator PK tests," in *Research in Parapsychology 1977.* Metuchen NJ: Scarecrow Press, 1978, pp. 157-162.

[87] H. Schmidt, "A PK test with electronic equipment," *J. Parapsychol.*, 34, pp. 175-181, 1970.

[88] H. Schmidt and L. Pantas, "Psi tests with internally different machines," *J. Parapsychol.*, 36, pp. 222-232, 1972.

[89] H. Schmidt, "PK effect on pre-recorded targets," *J. Amer. Soc. Psychical Research*, 70, pp. 267-291, 1976.

[90] ——, "PK tests with pre-recorded and pre-inspected seed numbers," in *Research in Parapsychology 1980.* Metuchen, NJ: Scarecrow Press, 1981, pp. 47-50.

[91] E. C. May, B. S. Humphrey, and G. S. Hubbard, "Electronic system perturbation techniques," SRI Int., Final Rep., Sept. 30, 1980.

[92] S. Krippner, Ed. *Advances in Parapsychological Research, Vol. 1, Psychokinesis.* New York: Plenum Press, 1977.

[93] B. Dunne, R. G. Jahn, and R. Nelson, "An REG experiment with large data-base capability," in *Research in Parapsychology 1981.* Metuchen, NJ: Scarecrow Press (in press).

[94] H. Puthoff and R. Targ, *Mind-Reach.* New York: Delacorte, 1977.

[95] ——, "Physics, entropy, and psychokinesis," in *Quantum Physics and Parapsychology*, L. Oteri, Ed. New York: Parapsychological Foundation, 1975.

[96] W. W. Eidson, D. L. Faust, K. Getsla, J. McClay, and B. Conoway, "An investigation of possible anomalous interaction between a human being and a polarized laser beam," preliminary rep. on investigations June-Sept. 1977, Drexel University, Physics Dep. (unpublished.)

[97] R. D. Nelson, B. J. Dunne, and R. G. Jahn, "Psychokinesis studies with a Fabry-Perot interferometer," in *Research in Parapsychology 1981.* Metuchen, NJ: Scarecrow Press (in press).

[98] J. B. Hasted, "Physical aspects of paranormal metal bending," *J. Soc. Psychical Research*, 49, pp. 583-607, 1977.

[99] ——, *The Metal-Benders.* London, Boston and Henley: Routledge and Kegan Paul, 1981.

[100] M. Ryzl, "A model for parapsychological communication," *J. Parapsychol.*, 30, pp. 18-31, 1966.

[101] M. A. Persinger, "Geophysical models for parapsychological experiences," *Psychoenergetic Syst.*, 1, pp. 63-74, 1975.

[102] G. E. Wolstenholme and E.C.P. Millar, Eds., *Extrasensory Perception: A CIBA Foundation Symposium.* New York: Citadel, 1966.

[103] O. C. de Beauregard, "Time symmetry and interpretation of quantum mechanics," *Foundations Phys.*, 6, pp. 539-559, 1976.

[104] G. Feinberg, "Possibility of faster-than light particles," *Phys. Rev.*, 159, p. 1089, 1967.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

[105] W. A. Tiller, "A lattice model of space and its relationship to multidimensional systems," in *Proc. A.R.E. Medical Symp.*, Phoenix, AZ, Jan. 1977.

[106] L. Oteri, Ed., *Quantum Physics and Parapsychology: Proceedings of an International Conference held in Geneva, Switzerland, August 26-27, 1974.* New York: Parapsychological Foundation, 1975.

[107] J.H.M. Whiteman, "Quantum theory and parapsychology," *J. Amer. Soc. for Psychical Research*, 67, pp. 341-360, 1973.

[108] E. H. Walker, "Consciousness and quantum theory," in *Psychic Exploration*, E. D. Mitchell *et al.*, J. White, Ed. New York: Putnam, 1974, pp. 544-568.

[109] R. D. Mattuck, "A quantum mechanical theory of the interaction between consciousness and matter," *Colloque International "Science et Conscience,"* Cordoba, Spain, Oct. 1-5, 1979.

[110] D. Bohm, "The enfolding-unfolding universe and consciousness," *Colloque International "Science et Conscience,"* Cordoba, Spain, Oct. 1-5, 1979.

[111] H. Margenau, "ESP in the framework of modern science," *J. Amer. Soc. Psychical Research*, 60, pp. 214-228, 1966.

[112] J.H.M. Whiteman, "Parapsychology and physics," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 730-756.

[113] A. Puharich, Ed., *The Iceland Papers: Select Papers on Experimental and Theoretical Research on the Physics of Consciousness.* Foreword by B. D. Josephson. Amherst, WI: Essentia Research Associates, 1979.

[114] A. H. Klopf, *Mental and Physical Phenomena: Toward a Unified Theory.* July 1977. (Unpublished manuscript.)

[115] Duke University, Dep. of Electrical Engineering; Princeton University, School of Engineering/Applied Science; Syracuse University, School of Computer and Information Science.

[116] S. S. Wilks, "Statistical aspects of experiments in telepathy," H.C.L.R. Lieber, Ed. A lecture by Samuel S. Wilks given at the Galois Institute of Mathematics, Long Island University, Mimeographed, 1938, 18 pp.

[117] P. Diaconis, "Statistical problems in ESP research," *Science*, 201, pp. 131-136, July 14, 1978.

[118] C.E.M. Hansel, *ESP and Parapsychology: A Critical Reevaluation.* Buffalo, NY: Prometheus Books, 1980.

[119] D. S. Burdick and E. F. Kelly, "Statistical Methods in Parapsychological Research," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, pp. 81-130, 1977.

[120] J. A. Greenwood and C. E. Stuart, "Mathematical techniques used in ESP research," *J. Parapsychol.*, 1, pp. 206-225, 1937.

[121] W. W. Carington, "Experiments on the paranormal cognition of drawings," *J. Parapsychol.*, 4, pp. 1-134, 1940.

[122] L. L. Gatlin, "A new measure of bias in finite sequences with applications to ESP data," *J. Amer. Soc. Psychical Research*, 1, pp. 29-43, 1979.

[123] C. T. Tart, "Randomicity, predictability, and mathematical inference strategies in ESP feedback experiments" (Discussion of Dr. Gatlin's paper). *J Amer. Soc. Psychical Research*, 1, pp. 44-60, 1979.

[124] J. G. Pratt, "Is ESP only a misnomer for response sequences chosen to match inferred target sequences?" (Discussion of Dr. Gatlin's paper.) *J. Amer. Soc. Psychical Research*, 1, pp. 60-66, 1979.

[125] D. S. Rogo, *Parapsychology: A Century of Inquiry.* New York: Taplinger, 1975, pp. 29-30.

[126] H. Bergson, "Presidential address to the Society for Psychical Research (1913)," *Proc. Soc. Psychical Research*, 27, pp. 157-175, 1914.

[127] ——, *Mind Energy.* New York: Henry Holt, 1920.

[128] C. D. Broad, *Religion, Philosophy and Psychical Research.* New York: Humanities Press, 1969.

[129] J. Ludwig, Ed., *Philosophy and Parapsychology.* Buffalo, NY: Prometheus Books, 1978.

[130] G. R. Schmeidler, Ed., *Parapsychology: Its Relation to Physics, Biology, Psychology, and Psychiatry.* Metuchen, NJ: Scarecrow Press, 1976.

[131] R. L. Van de Castle, "Parapsychology and anthropology," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 667-686.

[132] R. L. Morris, "Parapsychology, biology, and anpsi," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 687-715.

[133] W. H. Clark, "Parapsychology and religion," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 769-780.

[134] J. M. Backus, "Parapsychology and Literature," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 781-802.

[135] C. Panati, Ed., *The Geller Papers: Scientific Observations on the Paranormal Powers of Uri Geller.* Boston, MA: Houghton Mifflin, 1976.

[136] C. Crussard and J. Bouvaist, "Study of certain seemingly abnormal deformations and transformations of metals," *Memoires Scientifiques Revue Metallurgie*, Feb. 1978. (English translation H & P de Maigret, Nov. 1978).

[137] J. G. Pratt, *ESP Research Today: A Study of Developments in Parapsychology since 1960.* Metuchen, NJ: Scarecrow Press, 1973.

[138] J. G. Pratt, "The Pearisburg poltergeist," in *Research in Parapsychology 1977.* Metuchen, NJ: Scarecrow Press, 1978, pp. 174-182.

[139] J. Krieger, D. DalCorso, and G. Solfvin, "RSPK in Philadelphia," in *Research in Parapsychology 1978.* Metuchen, NJ: Scarecrow Press, 1979, pp. 59-60.

[140] A.R.G. Owen, *Can We Explain the Poltergeist?* New York: Garrett Helix, 1964.

[141] W. G. Roll, *The Poltergeist.* Metuchen, NJ: Scarecrow Press, 1976.

[142] W. G. Roll, "Poltergeists," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 382-413.

[143] W. Tiller, "A technical report on some psychoenergetic devices," *A.R.E. J.*, 7, p. 81, 1972.

[144] G. Schmeidler, "PK effects upon continuously recorded temperature," *J. Amer. Soc. Psychical Research*, 67, p. 325, 1973.

[145] R. G. Stanford, "Experimental psychokinesis: A review from diverse perspectives," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 324-381.

[146] H. Schmidt, "Psychokinesis," in *Psychic Exploration*, E. D. Mitchell *et al.*, J. White, Ed. New York: Putnam, 1974, pp. 179-193.

[147] W. Cox, "The effect of PK on the placement of falling objects," *J. Parapsychol.*, 15, pp. 40-48, 1951.

[148] ——, "The effect of PK on electromechanical systems," *J. Parapsychol.*, 29, pp. 165-175, 1965.

[149] J. Hasted, *The Metal Benders.* London, Boston and Henley: Routledge and Kegan Paul, 1981, p. 131.

[150] M. G. Shafer, "PK metal bending in a semi-formal small group," in *Research in Parapsychology 1980.* Metuchen, NJ: Scarecrow Press, 1981, pp. 32-35.

[151] C. Honorton and L. Tremmel, "Psitrek: A preliminary effort toward development of psi-conducive computer software," *Research in Parapsychology 1979.* Metuchen, NJ: Scarecrow Press, 1980, pp. 159-160.

[152] U. Sinclair, *Mental Radio* (Rev. 2nd printing, Introduction by W. McDougall, Preface by A. Einstein, with a report by W. F. Prince). Springfield, IL: Charles C. Thomas, 1962.

[153] H. E. Puthoff and R. Targ, "A perceptual channel for information transfer over kilometer distances: Historical perspective and recent research," in *Proc. IEEE*, vol. 64, pp. 329-354, 1976.

[154] R. Targ and H. E. Puthoff, "Information, transmission under conditions of sensory shielding," *Nature*, 252, pp. 602-607, 1974.

[155] C. T. Tart, H. E. Puthoff, and R. Targ, Eds., *Mind at Large: IEEE Symposia on the Nature of Extrasensory Perception.* New York: Praeger Special Studies, 1979.

[156] H. E. Puthoff, R. Targ, and E. C. May, "Experimental psi research: Implication for physics," in *The Role of Consciousness in the Physical World*, R. G. Jahn, Ed. Boulder, CO: Westview Press, 1981, pp. 37-86.

[157] B. J. Dunne and J. P. Bisaha, "Precognitive remote viewing in the Chicago area," *J. Parapsychol.*, 43, pp. 17-30, 1979.

[158] ——, "Multiple subject precognitive remote viewing," in *Research in Parapsychology 1977.* Metuchen, NJ: Scarecrow Press, 1978, pp. 146-151.

[159] ——, "Long distance precognitive remote viewing," in *Research in Parapsychology 1978.* Metuchen, NJ: Scarecrow Press, 1979, pp. 68-70.

[160] J. P. Bisaha and B. J. Dunne, "Long distance and multiple subject precognitive remote viewing of geographical locations," in *Mind at Large: IEEE Symposia on the Nature of Extrasensory Perception*, C. T. Tart, H. E. Puthoff, and R. Targ, Eds. New York: Praeger Special Studies, 1979, pp. 107-124.

[161] B. J. Dunne, *Precognitive Remote Perception.* M.A. thesis, Committee on Human Development, University of Chicago, 1979.

[162] B. J. Dunne and J. P. Bisaha, "Precognitive remote perception: A critical overview of the experimental program," in *Research in Parapsychology 1979.* Metuchen, NJ: Scarecrow Press, 1980, pp. 117-120.

[163] R. G. Jahn, "Psychic process, energy transfer, and things that go bump in the night," *Princeton Alumni Weekly*, Dec. 4, 1978, S-1-12.

[164] E. A. Rauscher, G. Weissmann, J. Sarfatti, and S.-P. Sirag, "Remote perception of natural scenes, shielded against ordinary perception," in *Research in Parapsychology 1975.* Metuchen,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

NJ: Scarecrow Press, pp. 41-45, 1976.

[165] S. Allen, P. Green, K. Rucker, R. Cohen, C. Goolsby, and R. L. Morris, "A remote viewing study using a modified version of the SRI procedure," in *Research in Parapsychology 1975*. Metuchen, NJ: Scarecrow Press, 1976, pp. 46-48.

[166] A. C. Hastings and D. B. Hurt, "A confirmatory remote viewing experiment in a group setting," *Proc. IEEE*, vol. 64, pp. 1544-1545, 1976.

[167] J. Vallee, A. C. Hastings, and G. Askevold, "Remote viewing through computer conferencing," *Proc. IEEE*, vol. 64, p. 1551, 1976.

[168] T. W. Whitson, D. N. Bogart, J. Palmer, and C. T. Tart, "Preliminary experiments in group remote viewing," *Proc. IEEE*, vol. 64, p. 1550, 1976.

[169] H. Chotas, "Remote viewing in the Durham area," *J. Parapsychol.*, 1, pp. 61-62, 1978.

[170] M. Schlitz and S. Deacon, "Remote viewing: A conceptual replication of Targ and Puthoff," in *Research in Parapsychology 1979*. Metuchen, NJ: Scarecrow Press, 1980, pp. 124-126.

[171] E. Karnes and E. Sussman, "Remote viewing: A response bias interpretation," *Psycholog. Rep.*, 44, pp. 471-479, 1979.

[172] E. Karnes, J. Ballou, E. Sussman, and F. Swaroff, "Remote viewing: Failures to replicate with control comparisons," *Psycholog. Rep.*, 45, pp. 963-973, 1979.

[173] E. W. Karnes, E. P. Sussman, P. Klusman, and L. Turcotte, "Failures to replicate remote viewing using psychic subjects," *Zetetic Scholar*, 6, pp. 66-76, 1980.

[174] M. Schlitz and E. Gruber, "Transcontinental remote viewing," *J. Parapsychol.*, 4, pp. 305-317, 1980.

[175] C. Scott, "On the evaluation of verbal material in parapsychology: A discussion of Dr. Pratt's monograph," *J. Soc. Psychical Research*, 46, pp. 79-90, 1972.

[176] R. L. Morris, "An exact method for evaluating preferentially matched free-response material," *J. Amer. Soc. Psychical Research*, 66, p. 401, 1972.

[177] G. Solfvin, E. Kelly, and D. Burdick, "Some new methods of analysis for preferential ranking data," *J. Amer. Soc. Psychical Research*, 72, p. 93, 1978.

[178] D. Marks and R. Kammann, *The Psychology of the Psychic*. Buffalo, NY: Prometheus Books, 1980.

[179] R. G. Jahn, B. J. Dunne, E. G. Jahn, "Analytical judging procedure for remote perception experiments," *J. Parapsychol.*, 3, pp. 207-231, 1980.

[180] B. J. Winer, *Statistical Principles in Experimental Design*. New York: McGraw-Hill, 1962, pp. 43-45.

[181] R. G. Jahn, Ed., *The Role of Consciousness in the Physical World*. Boulder, CO: Westview Press, 1981.

[182] R. G. Stanford, "An experimentally testable model for spontaneous psi events. I. Extrasensory events," *J. Amer. Soc. Psychical Research*, 68, pp. 34-57, 1974. "II. Psychokinetic events," *Ibid*, pp. 321-356.

[183] C.T.K. Chari, "Some generalized theories and models of psi: A critical evaluation," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 803-822.

[184] R. G. Stanford, "Conceptual frameworks of contemporary psi research," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 823-858.

[185] B. B. Wolman, "Mind and body: A contribution to a theory of parapsychological phenomena," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand Reinhold, 1977, pp. 861-879.

[186] A. M. Young, *The Geometry of Meaning*. New York: Delacorte Press/Seymour Lawrence, 1976.

[187] B. Shapin and L. Coly, Eds., *Concepts and Theories of Parapsychology: Proceedings of an International Conference held in New York, NY, Dec. 6, 1980*. New York: Parapsychology Foundation, Inc., 1981.

[188] F. Cazzamalli, "Phénomenes télépsychiques et radiations cerebrales," *Rev. Métapsych.*, 4, p. 3, 1925.

[189] I. M. Kogan, "Is telepathy possible?" *Radio Eng.*, 21, p. 75, 1966.
——, "Telepathy, hypotheses and observations," *Radio Eng.*, 22, p. 141, 1967.
——, "Information theory analysis of telepathic communication experiments," *Radio Eng.*, 23, p. 122, 1968.

[190] M. A. Persinger, "ELF field mediation in spontaneous psi events: Direct information transfer or conditional elicitation?" in *Mind at Large*, C. T. Tart, H. E. Puthoff, and R. Targ, Eds. New York: Praeger Special Studies, 1979, pp. 189-204.

[191] L. L. Gatlin, *Information Theory and the Living System*. New York: Columbia University Press, 1972.

[192] W. von Lucadou and K. Kornwachs, "Development of the system-theoretic approach to psychokinesis," paper presented at the Parascience Conf., London, 1979.

[193] K. Kornwachs and W. von Lucadou, "Psychokinesis and the concept of complexity," *Psychoenergetic Syst.*, 3, pp. 327-342, 1979.

[194] G. Feinberg, "Precognition—A memory of things future," in *Quantum Physics and Parapsychology*, L. Oteri, Ed. New York: Parapsychology Foundation, 1975, pp. 54-64.

[195] E. A. Rauscher, "Some physical models potentially applicable to remote perception," in *The Iceland Papers*, A. Puharich, Ed. Amherst, WI: Essentia Research Associates, 1979, pp. 49-83.

[196] A. S. Eddington, *Space, Time and Gravitation: An Outline of the General Theory of Relativity*. Cambridge, MA: Cambridge University Press, 1920, pp. 198-201.

[197] A. S. Eddington, *The Nature of the Physical World*. Cambridge, MA: Cambridge University Press, 1928, p. 276.

[198] K. Pribram, *Languages of the Brain*. Englewood Cliffs, NJ: Prentice-Hall, 1971.

[199] K. Pribram, "Toward a holonomic theory of perception," in *Gestalttheorie Modernen Psychologie*, S. Ertel, Ed. Darmstadt: Steinkopff, 1975.

[200] K. Pribram, "Problems concerning the structure of consciousness," in *Consciousness and the Brain*, G. Globus, Ed. New York: Plenum Press, 1976.

[201] D. Bohm, "Quantum theory as an indication of a new order in physics. Part A. The development of new orders as shown through the history of physics," *Foundations Phys.*, 1, pp. 359-381, 1971.

[202] D. Bohm, "Quantum theory as an indication of a new order in physics. Part B. Implicate and explicate order in physical law," *Foundations Phys.*, 1, pp. 139-168, 1971.

[203] A. Eddington, *Relativity Theory of Protons and Electrons*. New York: Macmillan, 1936, p. 329.

[204] O. C. de Beauregard, "The expanding paradigm of the Einstein paradox," in *The Iceland Papers*, A. Puharich, Ed. Amherst, WI: Essentia Associates, 1979, pp. 162-191.

[205] E. P. Wigner, "On hidden variables and quantum mechanical probabilities," *Amer. J. Phys.*, 38, p. 1005, 1970.

[206] J. S. Bell, "On the problem of hidden variables in quantum mechanics," *Rev. Modern Phys.*, 38, p. 447, 1966.

[207] D. Bohm and J. Bub, "A proposed solution of the measurement problem in quantum mechanics by a hidden variable theory," *Rev. Modern Phys.*, 38, p. 453, 1966.

[208] J. Mehra, "Quantum mechanics and the explanation of life," *Amer. Scientist*, 61, pp. 722-728, 1973.

[209] E. H. Walker, "Foundations of paraphysical and parapsychological phenomena," in *Quantum Physics and Parapsychology*, L. Oteri, Ed. New York: Parapsychology Foundation, 1975, pp. 1-44.

[210] R. G. Jahn and B. J. Dunne, "A quantum mechanical approach to the representation of paranormal phenomena," 1982 (to be published).

[211] T.N.E. Greville, "Are psi events random?" *J. Amer. Soc. Psychical Research*, 74, pp. 223-226, 1980.

[212] J. A. Greenwood and T.N.E. Greville, "On requirements for using statistical analysis in psi experiments," *J. Parapsychol.*, 43, pp. 315-321, 1979.

[213] R. Jahn and B. Dunne, "Physical models of psychic process," in *Concepts and Theories of Parapsychology*, B. Shapin and L. Coly, Eds. New York: Parapsychology Foundation, Inc., 1981, pp. 69-79.

[214] R. May, "Creativity and encounter," in *The Creative Imagination*, H. M. Ruitenbeek, Ed. Chicago, IL: Quadrangle Books, 1965, pp. 283-291.

[215] F. Barron, "The needs for order and disorder as motives in creative activity," in the second (1957) University of Utah research conference on the identification of creative scientific talent, C. W. Taylor, Ed. Salt Lake City: University of Utah Press, 1958, pp. 119-128.

[216] D. McClelland, "The calculated risk: An aspect of scientific performance," in *Scientific Creativity, its Recognition and Development*, C. W. Taylor and F. Barron, Eds. New York: Wiley, 1963, pp. 184-192.

[217] S. D. Drell, "When is a particle?" *Phys. Today*, 31, pp. 23-32, 1978.

[218] W. Heisenberg, "The nature of elementary particles," *Phys. Today*, 29, pp. 32-39, 1976.

[219] F. Capra, Paper LBL-796, Lawrence Berkeley Laboratories, University of California, Berkeley, CA, 1978.

[220] J. A. Wheeler, "The universe as home for man," *Amer. Scientist*, 62, pp. 683-691, 1974.

[221] W. W. Harman, "Broader implications of recent findings in psychological and psychic research," in *The Role of Consciousness in the Physical World*, R. G. Jahn, Ed. Boulder, CO: Westview Press, 1981, pp. 113-131.

[222] A. Angoff and B. Shapin, Eds., *A Century of Psychical Research: The Continuing Doubts and Affirmations*. New York: Parapsychology Foundation, 1971.

[223] C. A. Murchison, Ed. *The Case For and Against Psychical Belief*. New York: Arno Press, 1975. (Originally published in 1927.)

[224] W. F. Prince, *The Enchanted Boundary*. New York: Arno Press,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

1975. (Originally published in 1930.)

[225] K. M. Goldney and S. G. Soal, "Report on a series of experiments with Mrs. Eileen Garrett," *Proc. Soc. Psychical Research*, 45, pp. 43-87, 1938.

[226] J. Beloff, "The 'sweethearts' experiment," *J. Soc. Psychical Research*, 45, pp. 1-7, 1969.

[227] M. Maher, D. Peratsakis, and G. R. Schmeidler, "Cerebral localization effects in ESP processing: An attempted replication," *J Amer. Soc. for Psychical Research*, 73, pp. 167-177, 1979.

[228] S. Blackmore, "Correlation between ESP and memory," *Eur. J. Parapsychol.*, 3, pp. 127-147, 1980.

[229] J. Eisenbud, L. Hassel, H. Keely, and W. Sawrey, "A further study of teacher-pupil attitudes and results on clairvoyance tests in the fifth and sixth grades," *J. Amer. Soc. Psychical Research*, 54, pp. 72-80, 1960.

[230] E. Belvedere and D. Foulkes, "Telepathy and dreams: A failure to replicate," *Perceptual Motor Skills*, 33, pp. 783-789, 1971.

[231] J. Beloff and D. Bate, "An attempt to replicate the Schmidt findings," *J. Soc. Psychical Research*, 46, pp. 21-31, 1971.

[232] R. G. Jahn, reply to Bradley Dowden, *Zetetic Scholar*, 7, p. 127, 1980.

[233] C. C. Gillispie, Ed., *Dictionary of Scientific Biography*. New York: Scribner's, 1970-1980.

[234] R. L. Weber, *A Random Walk in Science*, E. Mendoza, Ed., Foreword by W. Cooper. New York: Crane, Russak, 1973, pp. 66-67.

[235] R. S. deRopp, *The New Prometheans*. New York: Delacorte, 1972, p. 2-28.

[236] L. Barnett, *The Universe and Dr. Einstein* (Rev. edit.). New York: The New American Library, 1952, p. 117.

[237] R. G. Jahn, "Psychic research: New dimensions or old delusions?" *Zetetic Scholar*, 6, pp. 5-16, 1980.

[238] J. Beloff, Ed., *New Directions in Parapsychology*. Metuchen, NJ: Scarecrow Press, 1975.

[239] J. B. Rhine and J. G. Pratt, *Parapsychology: Frontier Science of the Mind* (Rev. edit.). Springfield, IL: Charles C. Thomas, 1962.

[240] R. A. White, Ed., *Surveys in Parapsychology*. Metuchen, NJ: Scarecrow Press, 1976.

[241] R. A. White and L. A. Dale, *Parapsychology: Sources of Information*. Metuchen, NJ: Scarecrow Press, 1973. (Rev. edit. by R. A. White in preparation.)

[242] N. Bowles and F. Hynds with J. Maxwell, *Psi Search*. San Francisco, CA: Harper and Row, 1978.

[243] R. G. Jahn, B. J. Dunne, R. Nelson, and H. Mertz, "Anomalous operator influence on elementary microelectronic devices," Tech. note, School of Engineering/Applied Science, Princeton University, 1981.

[244] J. Bronowski, *Science and Human Values*. New York: Harper and Brothers, 1956.

[245] M. Polanyi, *Personal Knowledge*. Chicago, IL: University of Chicago Press, 1962.

[246] A. Korzybski, *Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics*. Clinton, MA: The Colonial Press, 1933.

[247] W. Penfield, *The Mystery of the Mind*. Princeton, NJ: Princeton University Press, 1975.

[248] M. Ferguson, *The Aquarian Conspiracy*. Los Angeles, CA: J. P. Tarcher, 1980.

[249] J. Salk, *Survival of the Wisest*. New York: Harper and Row, 1973.

[250] A. M. Young, *The Reflexive Universe*. New York: Delacorte Press, 1976.

[251] E. Jantsch, *The Self-Organizing Universe: Scientific and Human Implications of the Emerging Paradigms of Evolution*. Elmsford, NY: Pergamon Press, 1980.

[252] D. Dean *et al.*, *Executive ESP*. Englewood Cliffs, NJ: Prentice-Hall, 1974.

[253] P. T. de Chardin, *The Future of Man*. New York: Harper and Row, 1964 (pp. 166-167 in Harper Torchbooks edit.).

[254] Report to the House of Representatives, Committee on Science and Technology, June 1981.

[255] C. T. Tart, "A survey of expert opinion on potentially negative uses of psi, United States government interest in psi, and the level of research funding of the field," in *Research in Parapsychology 1978*. Metuchen, NJ: Scarecrow Press, 1979, pp. 54-55.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Psychology and Anomalous Observations

The Question of ESP in Dreams

Irvin L. Child Yale University

ABSTRACT: Books by psychologists purporting to offer critical reviews of research in parapsychology do not use the scientific standards of discourse prevalent in psychology. Experiments at Maimonides Medical Center on possible extrasensory perception (ESP) in dreams are used to illustrate this point. The experiments have received little or no mention in some reviews to which they are clearly pertinent. In others, they have been so severely distorted as to give an entirely erroneous impression of how they were conducted. Insofar as psychologists are guided by these reviews, they are prevented from gaining accurate information about research that, as surveys show, would be of wide interest to psychologists as well as to others.

In recent years, evidence has been accumulating for the occurrence of such anomalies as telepathy and psychokinesis, but the evidence is not totally convincing. The evidence has come largely from experiments by psychologists who have devoted their careers mainly to studying these anomalies, but members of other disciplines, including engineering and physics, have also taken part. Some psychologists not primarily concerned with parapsychology have taken time out from other professional concerns to explore such anomalies for themselves. Of these, some have joined in the experimentation (e.g., Crandall & Hite, 1983; Lowry, 1981; Radin, 1982). Some have critically reviewed portions of the evidence (e.g., Akers, 1984; Hyman, 1985). Some, doubting that the phenomena could be real, have explored nonrational processes that might encourage belief in their reality (e.g., Ayeroff & Abelson, 1976). Still others, considering the evidence substantial enough to justify a constructive theoretical effort, have struggled to relate the apparent anomalies to better established knowledge in a way that will render them less anomalous (e.g., Irwin, 1979) or not anomalous at all (e.g., Blackmore, 1984). These psychologists differ widely in their surmise about whether the apparent anomalies in question will eventually be judged real or illusory; but they appear to agree that the evidence to date warrants serious consideration.

Serious consideration of apparent anomalies seems an essential part of the procedures of science, regardless of whether it leads to an understanding of new discoveries or to an understanding of how persuasive illusions arise. Apparent anomalies—just like the more numerous observations that are not anomalous—can receive appropriate attention only as they become accurately known to the scientists to whose work they are relevant. Much parapsychological research is barred from being seriously considered because it is either neglected or misrepresented in writings by some psychologists—among them, some who have placed themselves in a prime position to mediate interaction between parapsychological research and the general body of psychological knowledge. In this article, I illustrate this important general point with a particular case, that of experimental research on possible ESP in dreams. It is a case of especially great interest but is not unrepresentative of how psychological publications have treated similar anomalies.

The Maimonides Research

The experimental evidence suggesting that dreams may actually be influenced by ESP comes almost entirely from a research program carried out at the Maimonides Medical Center in Brooklyn, New York. Among scientists active in parapsychology, this program is widely known and greatly respected. It has had a major indirect influence on the recent course of parapsychological research, although the great expense of dream-laboratory work has prevented it from being a direct model.

None of the Maimonides research was published in the journals that are the conventional media for psychology. (The only possible exception is that a summary of one study [Honorton, Krippner, & Ullman, 1972] appeared in convention proceedings of the American Psychological Association.) Much of it was published in the specialized journals of parapsychology. The rest was published in psychiatric or other medical journals, where it would not be noticed by many psychologists. Most of it was summarized in popularized form in a book (Ullman, Krippner, & Vaughan, 1973) in which two of the researchers were joined by a popular writer whose own writings are clearly not in the scientific tradition, and the book departs from the pattern of scientific reporting that characterizes the original research reports.

Copyright 1985 by the American Psychological Association, Inc. 0003-066X/85/$00.75
Vol. 40, No. 11, 1219-1230

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

How, then, would this research come to the attention of psychologists, so that its findings or its errors might in time be evaluated for their significance to the body of systematic observations upon which psychology has been and will be built? The experiments at Maimonides were published between about 1966 and 1972. In the years since—now over a decade—five books have been published by academic psychologists that purport to offer a scholarly review and evaluation of parapsychological research. They vary in the extent to which they seem addressed to psychologists themselves or to their students, but they seem to be the principal route by which either present or future psychologists, unless they have an already established interest strong enough to lead them to search out the original publications, might become acquainted with the experiments on ESP in dreams. I propose to review how these five books have presented knowledge about the experiments. First, however, I must offer a summary of the experiments; without that, my review would make sense only to readers already well acquainted with them.

The experiments at Maimonides grew out of Montague Ullman's observations, in his psychiatric practice, of apparent telepathy underlying the content of some dreams reported by his patients—observations parallel to those reported by many other psychiatrists. He sought to determine whether this apparent phenomenon would appear in a sleep laboratory under controlled conditions that would seem to exclude interpretations other than that of ESP. He was joined in this research by psychologist Stanley Krippner, now at the Saybrook Institute in San Francisco, and a little later by Charles Honorton, now head of the Psychophysical Research Laboratories in Princeton, New Jersey. Encouraged by early findings but seeking to improve experimental controls and identify optimal conditions, these researchers, assisted by numerous helpers and consultants, tried out various modifications of procedure. No one simple description of procedure, therefore, can be accurate for all of the experiments. But the brief description that follows is not, I believe, misleading as an account of what was generally done.

The Experimental Procedure

A subject would come to the laboratory to spend the night there as would-be percipient in a study of possible telepathic influence on dreams. He or she met and talked with the person who was going to serve as agent (that is, the person who would try to send a telepathic message), as well as with the two experimenters taking part that night, and procedures were explained in detail unless the percipient was a repeater for whom that step was not necessary. When ready to go to bed, the percipient was wired up in the usual way for monitoring of brain waves and eye movements, and he or she had no further contact with the agent or agent's experimenter until after the session was completed. The experimenter in the next room monitored the percipient's sleep and at the beginning of each period of rapid eye movements (REM), when it was reasonably certain the sleeper would be dreaming, notified the agent by pressing a buzzer.

The agent was in a remote room in the building, provided with a target picture (and sometimes accessory material echoing the theme of the picture) randomly chosen from a pool of potential targets as the message to be concentrated on. The procedure for random choice of a target from the pool was designed to prevent anyone else from knowing the identity of the target. The agent did not open the packet containing the target until isolated for the night (except for the one-way buzzer communication). Whenever signaled that the percipient had entered a REM period, the agent was to concentrate on the target, with the aim of communicating it telepathically to the percipient and thus influencing the dream the percipient was having. The percipient was oriented toward trying to receive this message. But of course if clairvoyance and telepathy are both possible, the percipient might have used the former—that is, might have been picking up information directly from the target picture, without the mediation of the agent's thoughts or efforts. For this reason, the term *general extrasensory perception (GESP)* would be used today, though the researchers more often used the term *telepathy*.

Toward the approximate end of each REM period, the percipient was awakened (by intercom) by the monitoring experimenter and described any dream just experienced (with prodding and questioning, if necessary, though the percipient of course knew in advance what to do on each awakening). At the end of the night's sleep, the percipient was interviewed and was asked for impressions about what the target might have been. (The interview was of course double-blind; neither percipient nor interviewer knew the identity of the target.) The dream descriptions and morning impressions and associations were recorded and later transcribed.

The original research reports and the popular book both present a number of very striking similarities between passages in the dream transcripts and the picture that happened to be the night's target. These similarities merit attention, yet they should in themselves yield no sense of conviction. Perhaps any transcript of a night's dreaming contains passages of striking similarity to any picture to which they might be compared. The Maimonides research, however, consisted of carefully planned experiments designed

Requests for reprints should be sent to Irvin L. Child at the Department of Psychology, Yale University, P.O. Box 11A, New Haven, Connecticut 06520-7447.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

to permit evaluation of this hypothesis of random similarity, and I must now turn to that aspect.

Results

To evaluate the chance hypothesis, the researchers obtained judgments of similarity between the dream content and the actual target for the night, and at the same time obtained judgments of similarity between the dream content and each of the other potential targets in the pool from which the target had been selected at random. The person judging, of course, had no information about which picture had been randomly selected as target; the entire pool (in duplicate) was presented together, with no clue as to which picture had been the target and which ones had not. That is, in the experimental condition a picture was randomly selected from a pool and concentrated on by the agent, and in the control condition a picture was left behind in the pool. Any consistent difference between target and nontarget in similarity to dream content, exceeding what could reasonably be ascribed to chance, was considered an apparent anomaly.

The data available for the largest number of sessions came from judgments made by judges who had no contact with the experiment except to receive (by mail, generally) the material necessary for judging (transcripts of dreams and interview and a copy of the target pool). For many sessions, judgments were also available from the dreamer; he or she, of course, made judgments only after completing participation in the experiment as dreamer (except in some series where a separate target pool was used for each night and the dreamer's judgments could be made at the end of the session). For many sessions, judgments were made for the dream transcripts alone and for the total transcript including the morning interview; for consistency I have used the latter, because it involved judges who had more nearly the same information as the subjects.

The only form in which the data are available for all series of sessions is a count of hits and misses. If the actual target was ranked in the upper half of the target pool, for similarity to the dreams and interview, the outcome was considered a hit. If the actual target was ranked in the lower half of the pool, the outcome was considered a miss. The hit-or-miss score is presented separately in Table 1 for judges and for subjects in the first two data columns. Where information is not supplied for one or the other, the reason is generally that it was impossible for the researchers to obtain it, and for a similar reason the number of cases sometimes varies.[1]

Each line of the table refers to one segment of the research, and segments for the most part are labeled as they were in the table of Ullman et al. (1973, pp. 275–277). Segments that followed the general procedure I described—all-night sessions, with an agent concentrating on the target during each of the percipient's REM periods—are gathered together in the first eight lines, A through H (in five of these segments, all but A, C, and H, a single percipient continued throughout a series, and in four of these the percipient was a psychologist). Other types of segments are presented in the rest of the table. Lines I, J, and K summarize precognitive sessions; here the target was not selected until after the dreaming and interview had been completed. The target consisted of a set of stimuli to be presented directly to the percipient after it had been selected in the morning. Lines L and M represent GESP sessions in which the percipient's dreams were monitored and recorded throughout the night, but the agent was attempting to transmit only before the percipient went to sleep or just after, or sporadically. Line N refers to a few clairvoyance sessions; these were like the standard GESP sessions except that there was no agent (no one knew the identity of the target). Finally, Line O reports on some GESP sessions in which each dream was considered separately; these formed a single experiment with four percipients, comparing nights involving a different target for each REM period with nights involving repeated use of a single target.

Regardless of the type of session (considering the five types I have described), each session fell into one of two categories: (a) pilot sessions, in which either a new dreamer or a new procedure was being tried out; these appear in lines H, K, and N, or (b) sessions in an experimental series, planned in advance as one or more sessions for each of two or more subjects, or as a number of sessions with the same dreamer throughout. Most of the researchers' publications were devoted to the results obtained in the experimental series, but the results of the pilot sessions have also been briefly reported.

A glance at the score columns for judges and for subjects is sufficient to indicate a strong tendency for an excess of hits over misses. If we average the outcome for judges and for subjects, we find that hits exceed misses on every one of the 15 independent lines on which outcome for hits and misses differs. (On Line E hits and misses occur with equal frequency.) By a simple sign-test, this outcome would be significant beyond the 0.0001 level. I would not stress the exact value here, for several reasons. There was no advance

[1] Of course, usable judgments could not be obtained from the subject in precognitive sessions, because at the time of judging he or she would already know what the target had been. For Line F, the single subject was unable to give the extra time required for judging, and for Line O one of the four subjects failed to make judgments. In a few of the pilot sessions (Lines H, K, and N) only the subject's judgment was sought, and in some sessions only that of one or more judges; in a few the mean judges' rating was neither a hit nor a miss but exactly at the middle.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 1
Summary of Maimonides Results on Tendency for Dreams to Be Judged More Like Target Than Like Nontargets in Target Pool

Series	Judges' score Hit	Judges' score Miss	Subjects' score Hit	Subjects' score Miss	z or t resulting from judgments: Judges	z or t resulting from judgments: Subjects	Sources
GESP: Dreams monitored and recorded throughout night; agent "transmitting" during each REM period							
A. 1st screening	7	5	10	2	$z = 0.71^b$	$z = \mathit{1.33}^b$	Ullman, Krippner, & Feldstein (1966)
B. 1st Erwin	5	2	6	1	$z = 2.53^b$	$z = 1.90^b$	Ullman et al. (1966)
C. 2nd screening	4	8	9	3	$z = -.25^b$	$z = \mathit{1.17}^b$	Ullman (1969)
D. Posin	6	2	6	2	$z = 1.05^c$	$z = 1.05^c$	Ullman (1969)
E. Grayeb	3	5	5	3	$z = -.63^c$	$z = 0.63^c$	Ullman, Krippner, & Vaughan (1973)
F. 2nd Erwin	8	0			$t = \mathit{4.93}^a$		Ullman & Krippner (1969)
G. Van de Castle	6	2	8	0	$t = 2.81^a$	$t = \mathit{2.74}^a$	Krippner & Ullman (1970)
H. Pilot sessions	53	14	42	22	$z = \mathit{4.20}^b$	$z = \mathit{2.21}^b$	Ullman et al. (1973)
Precognition: Dreams monitored and recorded throughout night; target experience next day							
I. 1st Bessent	7	1			$t = 2.81^a$		Krippner, Ullman, & Honorton (1971)
J. 2nd Bessent	7	1			$t = 2.27^a$		Krippner, Honorton, & Ullman (1972)
K. Pilot sessions	2	0			$z = 0.67^c$		Ullman et al. (1973)
GESP: Dreams monitored and recorded throughout night; agent active only at beginning or sporadically							
L. Sensory bombardment	8	0	4	4	$z = \mathit{3.11}^b$	$z = \mathit{0.00}^c$	Krippner, Honorton, Ullman, Masters, & Houston (1971)
M. Grateful Dead	7	5	8	4	$z = 0.61^c$	$z = 0.81^c$	Krippner, Honorton, & Ullman (1973)
Clairvoyance: Dreams monitored and recorded throughout night; concealed target known to no one							
N. Pilot sessions	5	3	4	5	$z = \mathit{0.98}^b$	$z = 0.00^b$	Ullman et al. (1973)
GESP: Single dreams							
O. Vaughan, Harris, Parise	105	98	74	79	$z = 0.63^c$	$z = -.32^c$	Honorton, Krippner, & Ullman (1972)

Note. GESP = general extrasensory perception. Italics identify results obtained with procedures that preserve independence of judgments in a series. For some series, the published source does not use the uniform measures entered in this table, and mimeographed laboratory reports were also consulted. Superscripts indicate which measure was available, in order of priority.
[a] Ratings. [b] Rankings. [c] Score (count of hits and misses).

plan to merge the outcomes for judges and subjects. Moreover, the various series could be split up in other ways. Although I think my organization of the table is very reasonable (and I did not notice this outcome until after the table was constructed), it is not the organization selected by Ullman et al. (1973); their table, if evaluated statistically in this same way, would not yield so striking a result. What is clear is that the tendency toward hits rather than misses cannot reasonably be ascribed to chance. There is some systematic—that is, nonrandom—source of anomalous resemblance of dreams to target.

Despite its breadth, this "hitting" tendency seems to vary greatly in strength. The data on single dreams—Line O—suggest no consistency. At the other extreme, some separate lines of the table look impressive. I will next consider how we may legitimately evaluate the relative statistical significance of

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

separate parts of the data on all-night sessions. (I will not try to take exact account here of the fact that the single-dream data are not significant, though it is wise to have in mind that the exact values I cite must be viewed as slightly exaggerated, in the absence of any explicit advance prediction that the results for all-night sessions and for single dreams would differ greatly.)

Two difficulties, one general and one specific, stand in the way of making as thorough an evaluation as I would wish. The general difficulty is that the researchers turned the task of statistical evaluation over to various consultants—for the most part, different consultants at various times—and some of the consultants must also have influenced the choice of procedures and measures. The consultants, and presumably the researchers themselves, seem not to have been at that time very experienced in working with some of the design problems posed by this research nor in planning how the research could be done to permit effective analysis. Much of the research was not properly analyzed at the time, and for much of it the full original data are no longer available. (The researchers have been very helpful in supplying me with material they have been able to locate despite dispersal and storage of the laboratory's files. Perhaps additional details may be recovered in the future.) The result is that completely satisfactory analysis is at present possible only for some portions of the data.

The specific difficulty results from a feature of the research design employed in most of the experimental series, a feature whose implications the researchers did not fully appreciate at the time. If a judge is presented with a set of transcripts and a set of targets and is asked to judge similarity of each target to each transcript, the various judgments may not be completely independent. If one transcript is so closely similar to a particular target that the judge is confident of having recognized a correct match, the judge (or percipient, of course) may minimize the similarity of that target to the transcripts judged later. Instructions to judges explicitly urged them to avoid this error, but we cannot tell how thoroughly this directive was followed. Nonindependence would create no bias toward either positive or negative evidence of correspondence between targets and transcripts, but it would alter variability and thus render inappropriate some standard tests of significance. I have entered in the two succeeding columns of the table a t or a z that can be used in evaluating the statistical significance of the departure from chance expectancy (t is required when ratings are available, and z must be used when only rankings or score counts are available, because sample variability in the former case is estimated from the data but in the latter case must be based conservatively on a theoretical distribution.) If ratings were available, they were used; if not, rankings were used if available; otherwise, score count was used.

Is there likely to have been much of this nonindependence in the series where it was possible? A pertinent fact is that the hits were not generally direct hits. That is, there was no overwhelming tendency for the correct target to be given first place rather than just being ranked in the upper half of the target pool. This greatly reduces the strength of the argument that ordinary significance tests are grossly inaccurate because of nonindependence. Because certainty is not possible, however, we need to separate results according to whether the procedures permitted this kind of nonindependence. In the table, I have italicized results that cannot have been influenced by this difficulty (either because each night's ratings were made by a different person or because each night in a series had, and was judged in relation to, a separate target pool) or that closely approximate this ideal condition.

The outcome is clear. Several segments of the data, considered separately, yield significant evidence that dreams (and associations to them) tended to resemble the picture chosen randomly as target more than they resembled other pictures in the pool. In the case of evaluation by outside judges, two of the three segments that are free of the problem of nonindependence yield separately significant results: The pilot sessions (Line H) yield a z of 4.20, and thus a p of .00002. An experiment with distant but multisensory targets (Line L) yields a z of 3.11 and a p of .001. If we consider segments in which judgments may not be completely independent of each other and analyze them in the standard way, we find that the two series with psychologist William Erwin as dreamer are also significant (if nonindependence of judgments does not seriously interfere), Line B with a z of 2.53 ($p < .01$) and Line F with a t of 4.93 and 7 df ($p < .01$). The two precognitive series (Lines I and J), each with 7 df, yield ts of 2.81 and 2.27, with p values slightly above and below .05, respectively.

Segment results based on the subjects' own judgments of similarity are less significant than those based on judgments by outside judges. Only two segments reach minimal levels of statistical significance: Line G, where the t of 2.74 with 7 df is significant at the .05 level, and Line H, where the z of 2.21 is significant at the .05 level.

The statistical evaluation of the separate segments of the Maimonides experiments also permits a more adequate evaluation of their overall statistical significance. For judgments by outside judges, three segments are free of the potential nonindependence of successive judgments (Lines H, L, and N). Putting these three together by the procedure Mosteller and Bush (1954, pp. 329–330) ascribed to Stouffer (recommended by Rosenthal [1984, p. 72] as the "simplest and most versatile" of the possible procedures), the joint p value is <.000002. For the subjects' own judgments, six segments are available (Lines A, C, G, H,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

L, and N), and their joint p value is less than .002. The other segments of the data have the problem of potential nonindependence of successive judgments, and even if the exaggeration of significance may be small for a single line, I would not want to risk compounding it in an overall p. Their prevailing unity of direction, however (direction not being subject to influence by the kind of nonindependence involved here), and the substantial size of some of the differences, justify the inference that the overall evidence of consistency far exceeds that indicated by only those selected segments for which a precise statistical statement is possible. The impression given by the mere count of hits and misses is thus fully confirmed when more sensitive measures are used.

Parapsychological experiments are sometimes criticized on the grounds that what evidence they provide for ESP indicates at most some very small effects detectable only by amassing large bodies of data. Those to whom this criticism has any appeal should be aware that the Maimonides experiments are clearly exempt from it. The significant results on Lines F and G of the table, for example, are each attributable basically to just eight data points.

If replications elsewhere should eventually confirm the statistically significant outcome of the Maimonides experiments, would the fact of statistical significance in itself establish the presence of the kind of anomaly called ESP? Of course not. Statistical significance indicates only the presence of consistency and does not identify its source. ESP, or the more general term *psi*, is a label for consistencies that have no identifiable source and that suggest transfer of information by channels not familiar to present scientific knowledge. A judgment about the appropriateness of the label, and thus about the "ESP hypothesis," is complex. It depends on a variety of other judgments and knowledge—how confidently other possible sources of the consistent effect can be excluded, whether other lines of experimentation are yielding results that suggest the same judgment, and so on.

I believe many psychologists would, like myself, consider the ESP hypothesis to merit serious consideration and continued research if they read the Maimonides reports for themselves and if they familiarized themselves with other recent and older lines of experimentation (e.g., Jahn, 1982, and many of the chapters in Wolman, 1977).

Some parapsychological researchers—among them the Maimonides group—have written at times as though a finding of statistical significance sufficiently justified a conclusion that the apparent anomaly should be classified as ESP. I can understand their choice of words, which is based on their own confidence that their experiments permitted exclusion of other interpretations. But perhaps psychologists who in the future become involved in this area may prefer to use a term such as *anomalies*, so as to avoid variable and possibly confusing connotations about the origin of the anomalies. Zusne and Jones (1982) wisely prepared the way for this usage in speaking of *anomalistic psychology*. But meanwhile, psychologists need not cut themselves off from knowledge of relevant facts because of dissatisfaction with the terminology surrounding their presentation.

Attempted Replications Elsewhere

The Maimonides pattern of controlled experiment in a sleep laboratory, obviously, is extremely time consuming and expensive, and replication seems to have been attempted so far at only two other sleep laboratories. At the University of Wyoming, two experiments yielded results approximately at mean chance expectation—slightly below in one study (Belvedere & Foulkes, 1971), slightly above in the other (Foulkes et al., 1972). In a replication at the Boston University School of Medicine (Globus, Knapp, Skinner, & Healey, 1968), overall results were not significantly positive, though in this instance encouragement for further exploration was reported. The researchers had decided in advance to base their conclusions on exact hits—that is, placing the target first, rather than just in the upper half; by this measure, the results were encouraging, though not statistically significant. Moreover, to quote the researchers, "*Post hoc* analysis revealed that the judges were significantly more correct when they were more 'confident' in their judgments. . . . Further conservatively designed research does seem indicated because of these findings" (Globus et al., 1968, p. 365).

A study by Calvin Hall (1967) is sometimes cited as a replication that confirmed the Maimonides findings; in truth, however, although it provided impressive case material, it was not done in a way that permits evaluation as a replication of the Maimonides experiments. Several small-scale studies, done without the facilities of a sleep laboratory, have been reported that are not replications of even one of the more ambitious Maimonides experiments but each of which reports positive results that might encourage further exploration (Braud, 1977; Child, Kanthamani, & Sweeney, 1977; Rechtschaffen, 1970; Strauch, 1970; Van de Castle, 1971). In the case of these minor studies—unlike the Maimonides studies and the three systematic replications—one must recognize the likelihood of selective publication on the basis of interesting results. Taken all together, these diverse and generally small-scale studies done elsewhere do, in my opinion, add something to the conviction the Maimonides experiments might inspire, that dream research is a promising technique for experimental study of the ESP question.

The lack of significant results in the three systematic replications is hardly conclusive evidence

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

against eventual replicability. In the Maimonides series, likewise, three successive replications (Lines C, D, and E in Table 1) yielded no significant result, yet they are part of a program yielding highly significant overall results.

If results of such potentially great interest and scientific importance as those of the Maimonides program had been reported on a more conventional topic, one might expect them to be widely and accurately described in reviews of the field to which they were relevant, and to be analyzed carefully as a basis for sound evaluation of whether replication and extension of the research were indicated, or of whether errors could be detected and understood. What has happened in this instance of anomalous research findings?

Representation of the Maimonides Research in Books by Psychologists

It is appropriate to begin with E. M. Hansel's 1980 revision of his earlier critical book on parapsychology. As part of his attempt to bring the earlier book up to date, he included an entire chapter on experiments on telepathy in dreams. One page was devoted to a description of the basic method used in the Maimonides experiments; one paragraph summarized the impressive outcome of 10 of the experiments. The rest of the chapter was devoted mainly to a specific account of the experiment in which psychologist Robert Van de Castle was the subject (the outcome is summarized in Line G of my Table 1) and to the attempted replication at the University of Wyoming (Belvedere & Foulkes, 1971), in which Van de Castle was again the subject. Another page was devoted to another of the Maimonides experiments that was also repeated at the University of Wyoming (Foulkes et al., 1972). Hansel did not mention the replication by Globus et al. (1968), whose authors felt that the results encouraged further exploration. Hansel gave more weight to the two negative outcomes at Wyoming than to the sum of the Maimonides research, arguing that sensory cues supposedly permitted by the procedures at Maimonides, not possible because of greater care taken by the Wyoming experimenters, were responsible for the difference in results. He did not provide, of course, the full account of procedures presented in the original Maimonides reports that might persuade many readers that Hansel's interpretation is far from compelling. Nor did he consider why some of the other experiments at Maimonides, not obviously distinguished in the care with which they were done from the two that were replicated (e.g., those on Lines E, M, and O of Table 1) yielded a close-to-chance outcome such as Hansel might have expected sensory cuing to prevent.

Hansel exaggerated the opportunities for sensory cuing—that is, for the percipient to obtain by ordinary sensory means some information about the target for the night. He did this notably by misinterpreting an ambiguous statement in the Maimonides reports, not mentioning that his interpretation was incompatible with other passages; his interpretation was in fact erroneous, as shown by Akers (1984, pp. 128–129). Furthermore, Hansel did not alert the reader to the great care exerted by the researchers to eliminate possible sources of sensory cuing. Most important is the fact that Hansel did not provide any plausible account—other than fraud—of how the opportunities for sensory cuing that he claimed existed would be likely to lead to the striking findings of the research. For example, he seemed to consider important the fact that at Maimonides the agent could leave his or her room during the night to go to the bathroom, whereas in Wyoming the agent had a room with its own bathroom, and the outer door to the room was sealed with tape to prevent the agent from emerging. Hansel did not attempt to say how the agent's visit to the bathroom could have altered the details of the percipient's dreams each night in a manner distinctively appropriate to that night's target. The only plausible route of influence on the dream record seems to be deliberate fraud involving the researchers and their subjects. The great number and variety of personnel in these studies—experimenters, agents, percipients, and judges—makes fraud especially unlikely as an explanation of the positive findings; but Hansel did not mention this important fact.

It appears to me that all of Hansel's criticisms of the Maimonides experiments are relevant only on the hypothesis of fraud (except for the mistaken criticism I have mentioned above). He said that unintentional communication was more likely but provided no evidence either that it occurred or that such communication—in any form in which it might have occurred—could have produced such consistent results as emerged from the Maimonides experiments. I infer that Hansel was merely avoiding making explicit his unsupported accusations of fraud. Fraud is an interpretation always important to keep in mind, and it is one that could not be entirely excluded even by precautions going beyond those used in the Wyoming studies. But the fact that fraud was as always, theoretically possible hardly justifies dismissal of a series of carefully conducted studies that offer important suggestions for opening up a new line of inquiry into a topic potentially of great significance. Especially regrettable is Hansel's description of various supposed defects in the experiments as though they mark the experiments as being carelessly conducted by general scientific criteria, whereas in fact the supposed defects are relevant only if one assumes fraud. A reader who is introduced to the Maimonides research by Hansel's chapter is likely to get a totally erroneous impression of the care taken by the experimenters to avoid various possible sources of error. The one thing they could

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

not avoid was obtaining results that Han[illegible] a priori impossible, hence evidence of fraud; but Hansel was not entirely frank about his reasoning.

An incidental point worth noting is that Hansel did not himself apply, in his critical attack, the standards of evidence he demanded of the researchers. His conclusions were based implicitly on the assumption that the difference of outcome between the Maimonides and the Wyoming experiments was a genuine difference, not attributable to random variation. He did not even raise the question, as he surely would have if, in some parallel instance, the Maimonides researchers had claimed or implied statistical significance where it was questionable. In fact, the difference of outcome might well have arisen from random error; for the percipient's own judgments the difference is significant at the 5% level (2-tailed), but for the outsiders' judgments it does not approach significance.

Another 1980 book is *The Psychology of Transcendence*, by Andrew Neher, in which almost 100 pages are devoted to "psychic experience." Neher differed from the other authors I refer to in describing the Maimonides work as a "series of studies of great interest" (p. 145), but this evaluation seems to be negated by his devoting only three lines to it and four lines to unsuccessful replications.

A third 1980 publication, *The Psychology of the Psychic*, by David Marks and Richard Kammann, provides less of a general review of recent parapsychology than Hansel's book or even Neher's one long chapter. It is largely devoted to the techniques of mentalists (that is, conjurors specializing in psychological rather than physical effects) and can be useful to anyone encountering a mentalist who pretends to be "psychic." Most readers are not likely to be aware that parapsychological research receives only limited attention. The jacket blurbs give a very different view of the book, as do the authors in their introductory sentences:

> ESP is just around the next corner. When you get there, it is just around the next corner. Having now turned over one hundred of these corners, we decided to call it quits and report our findings for public review. (Marks & Kammann, 1980, p. 4)

Given this introduction to the nature of the book, readers might suppose it would at least mention any corner that many parapsychologists have judged to be an impressive turning. But the Maimonides dream experiments received no mention at all.

Another volume, by psychologist James Alcock (1981), quite clearly purports to include a general review and evaluation of parapsychological research. Alcock mentioned (p. 6) that Hansel had examined the Maimonides experiments, but the only account of them that Alcock offered (on p. 163) was incidental to a discussion of control groups. By implication he seemed to reject [illegible] experiments because they included no control groups. He wrote that "a control group, for which no sender or no target was used, would appear essential" (p. 163). Later he added, "One could, alternatively, 'send' when the subject was not in the dream state, and compare 'success' in this case with success in dream state trials" (p. 163). The first of these statements suggests a relevant use of control groups but errs in calling it essential; in other psychological research, Alcock would have doubtless readily recognized that within-subject control can, where feasible, be much more efficient and pertinent than a separate control group. His second statement suggests a type of experiment that is probably impossible (because in satisfactory form it seems to require the subject to dream whether awake or asleep and not to know whether he or she was awake or asleep). This second kind of experiment, moreover, has special pertinence only to a comparison between dreaming and waking, not to the question of whether ESP is manifested in dreaming.

Alcock, in short, did not seem to recognize that the design of the Maimonides experiments was based on controls exactly parallel to those used by innumerable psychologists in other research with similar logical structure (and even implied, curiously enough, in his own second suggestion). He encouraged readers to think that the Maimonides studies are beyond the pale of acceptable experimental design, whereas in fact they are fine examples of appropriate use of within-subject control rather than between-subjects control.

The quality of thinking with which Alcock confronted the Maimonides research appeared also in a passage that did not refer to it by name. Referring to an article published in *The Humanist* by Ethel Grodzins Romm, he wrote,

> Romm (1977) argued that a fundamental problem with both the dream telepathy research and the remote viewing tests is that the reports suffer from what she called "shoe-fitting" language; she cited a study in which the sender was installed in a room draped in white fabric and had ice cubes poured down his back. A receiver who reported "white" was immediately judged to have made a "hit" by an independent panel. Yet, as she observed, words such as "miserable", "wet", or "icy" would have been better hits. . . . Again, the obvious need is for a control group. Why are they not used? (p. 163)

What Romm described as "shoe fitting" (misinterpreting events to fit one's expectations) is an important kind of error that is repeatedly made in interpretation of everyday occurrences by people who believe they are psychic. But the dream telepathy research at Maimonides was well protected against this kind of error by the painstaking controls that Alcock seemed not to have noticed. Surely Romm must be referring to some other and very sloppy dream research?

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Not at all. The details in this paragraph, and even more in Romm's article, point unmistakably, though inaccurately, to the fifth night of the first precognitive series at Maimonides. The actual details of target and response would alone deprive it of much of its value as an example of shoe fitting. As reported by Krippner, Ullman, & Honorton (1971), the target was a morning experience that included being in a room that was draped with white sheets. The subject's first dream report had included the statement, "I was just standing in a room, surrounded by white. Every imaginable thing in that room was white" (p. 201). There is more similarity here than Romm and Alcock acknowledged in mentioning from this passage only the single word "white."

More important, however, is the fact that the experiment they were referring to provided no opportunity for shoe fitting. The procedures followed in the experiment were completely misrepresented in a way that created the illusion that the possibility existed. There was no panel, in the sense of a group of people gathered together and capable of influencing each other. The judges, operating independently, separately judged every one of the 64 possible combinations of target and transcript yielded by the eight nights of the experiment, not just the eight correct pairings, and they had no clues to which those eight were. Their responses are hardly likely to have been immediate, as they required reading the entire night's transcript. Because each judge was working alone and was not recording times, there would have been no record if a particular response had been immediate, and no record of what particular element in the transcript led to an immediate response.

I looked up in a 1977 issue of *The Humanist* the article by Romm that Alcock cited. The half page on shoe-fitting language gave as examples this item from the Maimonides research and also the SRI remote-viewing experiments (Puthoff & Targ, 1976) done at SRI International. In both cases what was said was pure fiction, based on failure to note what was done in the experiments and in particular that the experimenters were well aware of the danger of shoe-fitting language and that the design of their experiments incorporated procedures to ensure that it could not occur. Romm's ignorance about the Maimonides research and her apparent willingness to fabricate falsehoods about it should be recognized by anyone who had read any of the Maimonides research publications. Yet Alcock accepted and repeated the fictions as though they were true. His presentation in the context of a book apparently in the scientific tradition seems to me more dangerous than Romm's original article, for anyone with a scientific orientation should be able to recognize Romm's article as propaganda. Its title, for example, is "When You Give a Closet Occultist a PhD, What Kind of Research Can You Expect?" and it repeatedly speaks of "cult phuds," meaning people with PhDs who are interested in parapsychological problems. Alcock's repetition of Romm's misstatements in a context lacking these clues may well be taken by many a reader as scholarly writing based on correct information and rational thought. Paradoxically, both Alcock's paragraph and Romm's article are excellent examples of the shoe-fitting error that both decry in others who are in fact carefully avoiding it.

The last of the five books that bring, or fail to bring, the Maimonides research to the attention of psychologists and their students is *Anomalistic Psychology: A Study of Extraordinary Phenomena of Behavior and Experience*, a 1982 volume by Leonard Zusne and Warren H. Jones. This is in many ways an excellent book, and it is also the one of the five that comes closest to including a general review of important recent research in parapsychology. Its brief account of the Maimonides dream experiments, however, misrepresented them in ways that should seriously reduce a reader's interest in considering them further.

Zusne and Jones's description of the basic procedure made three serious errors. First, it implied that one of the experimenters had a chance to know the identity of the target. ("After the subject falls asleep, an art reproduction is selected from a large collection randomly, placed in an envelope, and given to the agent" p. 260). In fact, precautions were taken to ensure that no one but the agent could know the identity of the target. Second, the authors stated that "three judges . . . rate their confidence that the dream content matches the target picture" (p. 260), leading the reader to suppose that the judges were informed of the identity of the target at the time of rating. In fact, a judge was presented with a dream transcript and a pool of potential targets and was asked to rate the degree of similarity between the transcript and each member of the pool, while being unaware of which member had been the target. Third, there was a similarly, though more obscurely, misleading description of how ratings were obtained from the dreamer.

This misinformation was followed by even more serious misrepresentation of the research and, by implication, of the competence of the researchers. Zusne and Jones (1982) wrote that Ullman and Krippner (1978) had found that dreamers were not influenced telepathically unless they knew in advance that an attempt would be made to influence them. This led, they wrote, to the subject's being "primed prior to going to sleep" through the experimenter's

> preparing the receiver through experiences that were related to the content of the picture to be telepathically transmitted during the night. Thus, when the picture was Van Gogh's Corridor of the St. Paul Hospital, which depicts a lonely figure in the hallways of a mental hospital, the receiver: (1)

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

heard Rozsa's *Spellbound* played on a phonograph; (2) heard the monitor laugh hysterically in the room; (3) was addressed as "Mr. Van Gogh" by the monitor; (4) was shown paintings done by mental patients; (5) was given a pill and a glass of water; and (6) was daubed with a piece of cotton dipped in acetone. The receiver was an English "sensitive," but it is obvious that no psychic sensitivity was required to figure out the general content of the picture and to produce an appropriate report, whether any dreams were actually seen or not. (pp. 260–261)

If researchers were to report positive results of the experiment described here by Zusne and Jones and were to claim that it provided some positive evidence of ESP, what would a reader conclude? Surely, that the researchers were completely incompetent, but probably not that they were dishonest. For dishonesty to take such a frank and transparent form is hardly credible.

Incompetence of the researchers is not, however, a proper inference. The simple fact, which anyone can easily verify, is that the account Zusne and Jones gave of the experiment is grossly inaccurate. What Zusne and Jones have done is to describe (for one specific night of the experiment) some of the stimuli provided to the dreamer the next morning, *after* his dreams had been recorded and his night's sleep was over. Zusne and Jones erroneously stated that these stimuli were provided *before* the night's sleep, to prime the subject to have or falsely report having the desired kind of dream. The correct sequence of events was quite clearly stated in the brief reference Zusne and Jones cited (Ullman & Krippner, 1978), as well as in the original research report (Krippner, Honorton, & Ullman, 1972).

I can understand and sympathize with Zusne and Jones's error. The experiment they cited is one in which the nocturnal dreamer was seeking to dream in response to a set of stimuli to be created and presented to him the next morning. As may be seen in Table 1, results from such precognitive sessions (all done with a single subject) were especially strong. This apparent transcendence of time as well as space makes the precognitive findings seem at least doubly impossible to most of us. An easy misreading, therefore, on initially scanning the research report, would be to suppose the stimuli to have been presented partly in advance (because some parts obviously involved a waking subject) and partly during sleep.

This erroneous reading on which Zusne and Jones based their account could easily have been corrected by a more careful rereading. In dealing with other topics, they might have realized the improbability that researchers could have been so grossly incompetent and could have checked the accuracy of their statements before publishing them. Zusne and Jones are not alone in this tendency to quick misperception of parapsychological research through preconception and prejudice; we have already seen it in Alcock's book. Alcock (1983) wrote the review of Zusne and Jones's book for *Contemporary Psychology*, the book-review journal of the American Psychological Association, and he did not mention this egregious error, even though very slight acquaintance with the Maimonides research should suffice to detect it.

Discussion

The experiments at the Maimonides Medical Center on the possibility of ESP in dreams clearly merit careful attention from psychologists who, for whatever reason, are interested in the question of ESP. To firm believers in the impossibility of ESP, they pose a challenge to skill in detecting experimental flaws or to the understanding of other sources of error. To those who can conceive that ESP might be possible, they convey suggestions about some of the conditions influencing its appearance or absence and about techniques for investigating it.

This attention is not likely to be given by psychologists whose knowledge about the experiments comes from the books by their fellow psychologists that purport to review parapsychological research. Some of those books engage in nearly incredible falsification of the facts about the experiments; others simply neglect them. I believe it is fair to say that none of these books has correctly identified any defect in the Maimonides experiments other than ones relevant only to the hypothesis of fraud or on inappropriate statistical reasoning (easily remedied by new calculations from the published data). I do not mean that the Maimonides experiments are models of design and execution. I have already called attention to a design flaw that prevents sensitive analysis of some of the experiments; and the control procedures were violated at one session, as Akers (1984) pointed out on the basis of the full information supplied in the original report. (Neither of these genuine defects was mentioned in any of the five books I have reviewed here, an indication of their authors' general lack of correct information about the Maimonides experiments.)

Readers who doubt that the falsification is as extreme as I have pictured it need only consult the sources I have referred to. Their doubt might also be reduced by familiarity with some of James Bradley's research (1981, 1984). In his 1984 article, he reported similar misrepresentations of fact on a topic, robustness of procedures of statistical inference, on which psychologists would not be thought to have nearly the strength of preconception that many are known to have about ESP. How much more likely, then, falsification on so emotionally laden a topic as ESP is for many psychologists! In the earlier article, Bradley (1981) presented experimental evidence (for college students, in this case, not psychologists) that confi-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

dence in the correctness of one's own erroneous opinions is positively correlated with the degree of expertise one believes oneself to have in the field of knowledge within which the erroneous opinion falls. This finding may help in understanding why the authors of some of these books did not find it necessary to consider critically their own erroneous statements.

A very considerable proportion of psychologists have a potential interest in the question of ESP. In a recent survey (Wagner & Monnet, 1979) of university professors in various fields, 34% of psychologists were found to consider ESP either an established fact or a likely possibility, exactly the same proportion as considered it an impossibility. In this survey, psychologists less frequently expressed a positive opinion than did members of other disciplines, a finding that may be attributable to psychologists' better understanding of sources of error in human judgment. There seems to be no equally sound reason for the curious fact that psychologists differed overwhelmingly from others in their tendency to consider ESP an impossibility. Of natural scientists, only 3% checked that opinion; of the 166 professors in other social sciences, not a single one did.

Both of these groups of psychologists have been ill served by the apparently scholarly books that seem to convey information about the dream experiments. The same may be said about some other lines of parapsychological research. Interested readers might well consult the original sources and form their own judgments.

REFERENCES

Akers, C. (1984). Methodological criticisms of parapsychology. In S. Krippner (Ed.), *Advances in parapsychological research* (Vol. 4, pp. 112–164). Jefferson, NC: McFarland.

Alcock, J. E. (1981). *Parapsychology, science or magic? A psychological perspective.* New York: Pergamon Press.

Alcock, J. E. (1983). Bringing anomalies back into psychology. *Contemporary Psychology, 28,* 351–352.

Ayeroff, F., & Abelson, R. P. (1976). ESP and ESB: Belief in personal success at mental telepathy. *Journal of Personality and Social Psychology, 34,* 240–247.

Belvedere, E., & Foulkes, D. (1971). Telepathy and dreams: A failure to replicate. *Perceptual and Motor Skills, 33,* 783–789.

Blackmore, S. J. (1984). A psychological theory of the out-of-body experience. *Journal of Parapsychology, 48,* 201–218.

Bradley, J. V. (1981). Overconfidence in ignorant experts. *Bulletin of the Psychonomic Society, 17,* 82–84.

Bradley, J. V. (1984). Antinonrobustness: A case study in the sociology of science. *Bulletin of the Psychonomic Society, 22,* 463–466.

Braud, W. (1977). Long-distance dream and presleep telepathy. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in parapsychology 1976* (pp. 154–155). Metuchen, NJ: Scarecrow.

Child, I. L., Kanthamani, H., & Sweeney, V. M. (1977). A simplified experiment in dream telepathy. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in parapsychology 1976* (pp. 91–93). Metuchen, NJ: Scarecrow.

Crandall, J. E., & Hite, D. D. (1983). Psi-missing and displacement: Evidence for improperly focused psi? *Journal of the American Society for Psychical Research, 77,* 209–228.

Foulkes, D., Belvedere, E., Masters, R. E. L., Houston, J., Krippner, S., Honorton, C., & Ullman, M. (1972). Long-distance "sensory-bombardment" ESP in dreams: A failure to replicate. *Perceptual and Motor Skills, 35,* 731–734.

Globus, G., Knapp, P., Skinner, J., & Healey, J. (1968). An appraisal of telepathic communication in dreams. *Psychophysiology, 4,* 365.

Hall, C. (1967). Experimente zur telepathischen Beeinflussung von Träumen. [Experiments on telepathic influence on dreams]. *Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 10,* 18–47.

Hansel, C. E. M. (1980). *ESP and parapsychology: A critical reevaluation.* Buffalo, NY: Prometheus.

Honorton, C., Krippner, S., & Ullman, M. (1972). Telepathic perception of art prints under two conditions. *Proceedings of the 80th Annual Convention of the American Psychological Association, 7,* 319–320.

Hyman, R. (1985). The ganzfeld psi experiment: A critical appraisal. *Journal of Parapsychology, 49,* 3–49.

Irwin, H. J. (1979). *Psi and the mind: An information processing approach.* Metuchen, NJ: Scarecrow.

Jahn, R. G. (1982). The persistent paradox of psychic phenomena: An engineering perspective. *Proceedings of the Institute of Electrical and Electronics Engineers, 70,* 136–170.

Krippner, S., Honorton, E., & Ullman, M. (1972). A second precognitive dream study with Malcolm Bessent. *Journal of the American Society for Psychical Research, 66,* 269–279.

Krippner, S., Honorton, C., & Ullman, M. (1973). An experiment in dream telepathy with "The Grateful Dead." *Journal of the American Society of Psychosomatic Dentistry and Medicine, 20,* 9–17.

Krippner, S., Honorton, C., Ullman, M., Masters, R., & Houston, J. (1971). A long-distance "sensory-bombardment" study of ESP in dreams. *Journal of the American Society for Psychical Research, 65,* 468–475.

Krippner, S., & Ullman, M. (1970). Telepathy and dreams: A controlled experiment with electroencephalogram-electro-oculogram monitoring. *Journal of Nervous and Mental Disease, 151,* 394–403.

Krippner, S., Ullman, M., & Honorton, C. (1971). A precognitive dream study with a single subject. *Journal of the American Society for Psychical Research, 65,* 192–203.

Lowry, R. (1981). Apparent PK effect on computer-generated random digit series. *Journal of the American Society for Psychical Research, 75,* 209–220.

Marks, D., & Kammann, R. (1980). *The psychology of the psychic.* Buffalo, NY: Prometheus.

Mosteller, F., & Bush, R. R. (1954). Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology* (Vol. 1, pp. 289–334). Cambridge, MA: Addison-Wesley.

Neher, A. (1980). *The psychology of transcendence.* Englewood Cliffs, NJ: Prentice-Hall.

Puthoff, H. E., & Targ, R. (1976). A perceptual channel for information transfer over kilometer distances: Historical perspective and recent research. *Proceedings of the Institute of Electrical and Electronic Engineers, 64,* 329–354.

Radin, D. I. (1982). Experimental attempts to influence pseudorandom number sequences. *Journal of the American Society for Psychical Research, 76,* 359–374.

Rechtschaffen, A. (1970). Sleep and dream states: An experimental design. In R. Cavanna (Ed.), *Psi favorable states of consciousness* (pp. 87–120). New York: Parapsychology Foundation.

Romm, E. G. (1977). When you give a closet occultist a Ph.D., what kind of research can you expect? *The Humanist, 37*(3), 12–15.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage.

Strauch, I. (1970). Dreams and psi in the laboratory. In R. Cavanna

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

(Ed.), *Psi favorable states of consciousness* (pp. 46–54). New York: Parapsychology Foundation.

Ullman, M. (1969). Telepathy and dreams. *Experimental Medicine & Surgery, 27,* 19-38.

Ullman, M., & Krippner, S. (1969). A laboratory approach to the nocturnal dimension of paranormal experience: Report of a confirmatory study using the REM monitoring technique. *Biological Psychiatry, 1,* 259-270.

Ullman, M., & Krippner, S. (1978). Experimental dream studies. In M. Ebon (Ed.), *The Signet handbook of parapsychology* (pp. 409–422). New York: New American Library.

Ullman, M., Krippner, S., & Feldstein, S. (1966). Experimentally induced telepathic dreams: Two studies using EEG-REM monitoring technique. *International Journal of Neuropsychiatry, 2,* 420–437.

Ullman, M., Krippner, S., & Vaughan, A. (1973). *Dream telepathy.* New York: Macmillan.

Van de Castle, R. L. (1971). The study of GESP in a group setting by means of dreams. *Journal of Parapsychology, 35,* 312.

Wagner, M. W., & Monnet, M. (1979). Attitudes of college professors toward extra-sensory perception. *Zetetic Scholar,* no. 5, 7-16.

Wolman, B. B. (Ed.). (1977). *Handbook of parapsychology.* New York: Van Nostrand Reinhold.

Zusne, L., & Jones, W. H. (1982). *Anomalistic psychology: A study of extraordinary phenomena of behavior and experience.* Hillsdale, NJ: Erlbaum.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Introduction to PSYCHOLOGY

TENTH EDITION

Rita L. Atkinson
University of California, San Diego

Richard C. Atkinson
University of California, San Diego

Edward E. Smith
University of Michigan

Daryl J. Bem
Cornell University

HARCOURT BRACE JOVANOVICH, PUBLISHERS
San Diego New York Chicago Austin Washington, D.C.
London Sydney Tokyo Toronto

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

difference between sleep and hypnosis in the EEG and by demonstrations of alert hypnosis.

A psychoanalytic theory suggests that hypnosis is a state of *partial regression* in which the subject lacks the controls present in normal waking consciousness and therefore acts impulsively and engages in fantasy production. The idea is that hypnosis causes a regression in the thought processes to a more infantile stage; fantasies and hallucinations during hypnosis are indicators of a primitive mode of thought uncensored by higher levels of control (Gill, 1972).

A theory based on the dramatic nature of many hypnotic behaviors emphasizes a kind of involuntary *role enactment* as a response to social demands. This theory does *not* imply that the subject is playacting in a deliberate attempt to fool the hypnotist; it assumes that the subject becomes so deeply involved in a role that actions take place without conscious intent (Coe & Sarbin, 1977).

Yet another approach emphasizes the dissociative aspects of hypnosis. Dissociation involves a split of consciousness into several streams of thought, each somewhat independent of the others. Hypnosis theoretically induces a dissociative state in the subject so that he is not aware of all that is occurring in consciousness. The hypnotist, however, can tap into the various streams of thought. A special version of this theory, called *neodissociation theory*, has proved to be useful in analyzing hypnotic phenomena (see the following Critical Discussion).

Competing theories of hypnosis were argued more vehemently in the 1960s and 1970s than they are today. With the facts and relationships now better understood, differences between explanations fade in importance. Each theory calls attention to some significant features of hypnosis, and as new data become available, differences are being resolved (Kihlstrom, 1987).

PSI PHENOMENA

A discussion of consciousness would not be complete without considering some esoteric and mystical claims about the mind that have attracted widespread public attention. Of particular interest are questions about whether or not human beings can a) acquire information about the world or other people in ways that do not involve stimulation of the known sense organs, or b) influence physical events by purely mental means. These questions are the source of controversy over the existence of *psi*, processes of information and/or energy exchange not currently explicable in terms of known science (in other words, known physical mechanisms). The phenomena of psi are the subject matter of *parapsychology* ("beside psychology") and include the following:

1. *Extrasensory perception* (ESP). Response to external stimuli without any known sensory contact.
   a. *Telepathy*. Thought transference from one person to another without the mediation of any known channel of sensory communication (for example, identifying a playing card merely being thought of by another person)
   b. *Clairvoyance*. Perception of objects or events that do not provide a stimulus to the known senses (for example, identifying a concealed playing card whose identity is unknown to anyone).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

c. *Precognition*. Perception of a future event that could not be anticipated through any known inferential process (for example, predicting that a particular number will come up on the next throw of dice)

2. *Psychokinesis* (PK). Mental influence over physical events without the intervention of any known physical force (for example, willing that a particular number will come up on the throw of dice).

Experimental Evidence

Most parapsychologists consider themselves to be scientists applying the usual rules of scientific inquiry to admittedly unusual phenomena. Yet the phenomena of psi are so extraordinary and so similar to what are widely regarded as superstitions that some scientists declare psi to be an impossibility and reject the legitimacy of parapsychological inquiry. Such a priori judgments are out of place in science; the real question is whether the empirical evidence is acceptable by scientific standards. Many psychologists who are not yet convinced that psi has been demonstrated are nevertheless open to the possibility that new evidence might emerge that would be more compelling. For their part, many parapsychologists believe that several recent experimental procedures either provide that evidence already or hold the potential for doing so. We shall examine one of the most promising of these, the *ganzfeld procedure*.

The ganzfeld procedure tests for telepathic communication between a subject acting as the "receiver" and another subject who serves as the "sender." The receiver is sequestered in an acoustically isolated room and placed in a mild form of perceptual isolation: translucent ping-pong ball halves are taped over the eyes and headphones are placed over the ears; diffuse red light illuminates the room, and white noise is played through the headphones. (White noise is a random mixture of sound frequencies similar to the hiss made by a radio tuned between stations.) This homogeneous visual and auditory environment is called the *Ganzfeld*, a German word meaning "total field."

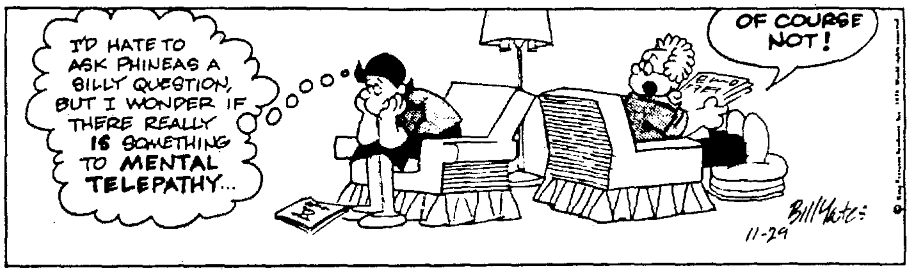
The sender is sequestered in a separate acoustically isolated room, and a visual stimulus (picture, slide, or brief videotape sequence) is randomly selected from a large pool of similar stimuli to serve as the "target" for the session. While the sender concentrates on the target, the receiver attempts to describe it by providing a continuous verbal report of his or her ongoing imagery and free associations. Upon completion of the session, the receiver is presented with four stimuli—one of which is the target—and asked to rate the degree to which each matches the imagery and associations experienced during the ganzfeld session. A "direct hit" is scored if the receiver assigns the highest rating to the target stimulus.

More than 50 experiments have been conducted since the procedure was first introduced in 1974; the typical experiment involves about 30 ganzfeld sessions in which a receiver attempts to identify the target transmitted by the sender. An overall analysis of 28 studies (comprising a total of 835 ganzfeld sessions conducted by investigators in 10 different laboratories) reveals that subjects were able to select the correct target stimulus 38 percent of the time. Because a subject must select the target from four alternatives, we would expect a success rate of 25 percent if only chance were operating. Statistically this result is highly significant; the probability that it could have arisen by chance is less than one in a billion (Honorton, 1985).

The receiver (top photograph) and the sender (bottom photograph) in a Ganzfeld experiment.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4



© Bill Yates; reprinted with special permission of King Features Syndicate, Inc., 1971.

Debate over the Evidence

In 1985 and 1986, the *Journal of Parapsychology* published an extended examination of the ganzfeld studies, focusing on a debate between Ray Hyman, a cognitive psychologist and critic of parapsychology, and Charles Honorton, a parapsychologist and major contributor to the ganzfeld database. They agree on the basic quantitative results but disagree on points of interpretation (Hyman, 1985; Hyman & Honorton, 1986; Honorton, 1985). We shall use their debate as a vehicle for examining the issues involved in evaluating claims of psi.

REPLICATION PROBLEM In science generally, a phenomenon is not considered established until it has been observed repeatedly by several researchers. Accordingly, the most serious criticism of parapsychology is that it has failed to produce a single reliable demonstration of psi that can be replicated by other investigators. Even the same investigator testing the same individuals over time may obtain statistically significant results on one occasion but not on another. The ganzfeld procedure is no exception; fewer than half (43 percent) of the 28 studies analyzed in the debate yielded statistically significant results.

The parapsychologists' most effective response to this criticism actually comes from within psychology itself. Many statisticians and psychologists are dissatisfied with psychology's focus on the *statistical significance level* as the sole measure of a study's success. As an alternative, they are increasingly adopting the technique of *meta-analysis*, a statistical technique that treats the accumulated studies of a particular phenomenon as a single grand experiment and each study as a single observation. Thus any study that obtains results in the positive direction—even though it may not be statistically significant itself—contributes to the overall strength and reliability of the phenomenon rather than simply being dismissed as a failure to replicate (Glass, McGaw, & Smith, 1981; Rosenthal, 1984).

From this perspective, the ganzfeld studies provide impressive replicability: 23 of the 28 studies obtain positive results (more direct hits than chance would predict), a result whose probability of occurring by chance is less than one in a thousand.

The conventional criterion of replication further requires that any competent investigator be able to reproduce the claimed phenomenon, not just one or two gifted experimenters. This is often a difficult criterion to achieve in new areas of investigation because a number of unsuspected variables might affect the outcome. In psychological experiments, the experimenter is often an important social stimulus for the subject and hence a poorly

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

controlled source of variability. Even in such an established area as classical conditioning, investigators at one university were obtaining positive results 94 percent of the time while other investigators could do so only 62 percent of the time (Rosenthal, 1966; Spence, 1964). Nor is the field of psychology alone here. Similar replication difficulties have been reported in medical studies of placebo efficacy (Moerman, 1981) and in such physical science areas as laser technology (Collins, 1974).

This problem may be even more acute in parapsychology because psi effects may legitimately depend on the motivational atmosphere established by the experimenter. Some parapsychologists further believe that the experimenter's own psi abilities and attitudes can have an effect.

Despite these potential difficulties, the replicability of the ganzfeld effect does not appear to rest on the success of one or two investigators. Six of the 10 investigators contributing to the 28 examined studies obtained statistically significant results; and, even if all the studies of the two most successful investigators are discarded from the analysis (half of the studies), the results remain significant (Palmer, Honorton, & Utts, 1988).

The power of a particular experiment to replicate an effect also depends on how strong the effect is and how many observations are made. If an effect is weak, an experiment with too few subjects or observations will fail to detect it at a statistically significant level—even though the effect actually exists.

This is strikingly illustrated by a recent medical experiment designed to determine whether aspirin can prevent heart attacks. The study was discontinued in 1987 because it was already clear the answer was yes. After six years, the aspirin group had already suffered 45 percent fewer heart attacks than a control group that received only placebo medication, a result that would occur by chance less than one time out of a million (The Steering Committee of the Physicians' Health Study Research Group, 1988). With such impressive results, it was considered unethical to keep the control group on placebo medication. This study was widely publicized as a major medical breakthrough.

The pertinent point here is that the study included over 22,000 subjects. If it were to be repeated with 3,000 subjects, a significant aspirin effect would be unlikely to emerge; the experiment would fail to replicate. Despite its undisputed reality and its practical importance, the aspirin effect is actually quite weak.

Now reconsider the ganzfeld effect. If the effect actually exists and has a true direct-hit rate of 38 percent, then statistically we should expect studies with 30 ganzfeld sessions (the average for the 28 studies) to obtain a statistically significant psi effect only about one-third of the time (Utts, 1986). The ganzfeld effect is about four times stronger than the aspirin effect.

In short, it is unrealistic to demand that any real effect be replicable at any time by any competent investigator. The replication issue is more complex than that, and meta-analysis is proving to be a valuable tool for dealing with some of those complexities.

INADEQUATE CONTROLS The second major criticism of parapsychology is that many, if not most, of the experiments have inadequate controls and safeguards. Flawed procedures that would permit a subject to obtain the communicated information in normal sensory fashion either inadvertently, or through deliberate cheating, are particularly fatal. This is called the problem of *sensory leakage*. Inadequate procedures for randomizing (randomly selecting) target stimuli are another common problem.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Methodological inadequacies plague all the sciences, but the history of parapsychology is embarrassingly full of promising results that collapsed when the procedures were critically examined (Akers, 1984). One common charge against parapsychology is that preliminary, poorly controlled studies often obtain positive results but that as soon as better controls and safeguards are introduced, the results disappear.

Once a flaw is discovered in a completed experiment, there is no persuasive way of arguing that the flaw did not contribute illegitimately to a positive outcome; the only remedy is to redo the experiment correctly. In a database of several studies, however, meta-analysis can evaluate the criticism empirically by checking to see if, in fact, the more poorly controlled studies obtained more positive results than did the better controlled studies. If there is a correlation between a procedural flaw and positive results across the studies, then there is a problem. In the case of the ganzfeld database, both critic Hyman and parapsychologist Honorton agree that flaws of inadequate security and possible sensory leakage do not correlate with positive results. Hyman claimed to find a correlation between flaws of randomization and positive results, but both Honorton's analysis and two additional analyses by nonparapsychologists dispute his conclusion (Harris & Rosenthal, 1988; Saunders, 1985). Moreover, a series of 10 new studies designed to control for flaws identified in the original database yielded results consistent with the original set of 28 studies (Harris & Rosenthal, 1988).

FILE-DRAWER PROBLEM Suppose that each of 20 investigators independently decides to conduct a ganzfeld study. Even if there were no genuine ganzfeld effect, there is a reasonable probability that at least one of these investigators would obtain a statistically significant result by pure chance. That lucky investigator would then publish a report of the experiment, but the other 19 investigators—all of whom obtained *null* results—are likely to become discouraged, put their data into a file drawer, and move onto something more promising. As a result, the scientific community would learn about the one successful study but have no knowledge of the 19 null studies buried away in the file drawers. The database of known studies would thus be seriously biased toward positive studies, and any meta-analysis of that database would arrive at similarly biased conclusions. This is called the *file-drawer problem*.

The problem is particularly tricky because it is impossible, by definition, to know how many unknown studies are languishing in file drawers somewhere. Nevertheless, parapsychologists offer two lines of defense against the charge that the file-drawer problem seriously compromises their database.

First, they point out that the *Journal of Parapsychology* actively solicits and publishes studies that report negative findings. Moreover, the community of parapsychologists is relatively small, and most investigators are cognizant of ongoing work in the various laboratories around the world. When conducting meta-analysis, parapsychologists actively attempt to scout out unpublished negative studies at conventions and through their personal networks.

But their major defense is statistical, and again meta-analysis provides an empirical approach to the problem. By knowing the overall statistical significance of the known database, it is possible to compute the number of studies with null results that would have to exist in file drawers to cancel out that significance. In the case of the ganzfeld database, there would have to be

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

over 400 unreported studies with null results—the equivalent of 12,000 ganzfeld sessions—to cancel out the statistical significance of the 28 studies analyzed in the debate (Honorton, 1985). Not surprisingly, there is consensus that the overall significance of the ganzfeld studies cannot reasonably be explained by the file-drawer effect (Hyman & Honorton, 1986).

Rather than continuing their debate, Hyman and Honorton issued a joint communiqué in which they set forth their areas of agreement and disagreement and made a series of suggestions for the conduct of future ganzfeld studies (Hyman & Honorton, 1986). Their debate and the subsequent discussion provide a valuable model for evaluating disputed domains of scientific inquiry.

Anecdotal Evidence

In the public's mind, the evidence for psi consists primarily of personal experiences and anecdotes. Such evidence is unpersuasive in science because it suffers fatally from the same problems that jeopardize the experimental evidence—nonreplicability, inadequate controls, and the file-drawer problem.

The replication problem is acute because most such evidence consists of one-time occurrences. A woman announces a premonition that she will win the lottery that day—and she does. You dream about an unlikely event that actually occurs a few days later. A "psychic" correctly predicts the assassination of a public figure. Such incidents may be subjectively compelling, but there is no way to evaluate them because they are not repeatable.

The problem of inadequate controls and safeguards is decisive because such incidents occur under unexpected and ambiguously specified conditions. There is thus no way of ruling out such alternative interpretations as coincidence (chance), faulty memories, and deliberate deception.

A psychic at work in New Orleans.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

And finally, the file-drawer problem is also fatal. The lottery winner who announced ahead of time that she would win is prominently featured in the news. But the thousands of others with similar premonitions who did *not* win are never heard from; they remain in the file drawers. It is true that the probability of this woman's winning the lottery was very low. But the critical criterion in evaluating this case is not the probability that *she* would win but the probability that any *one* of the thousands who thought they would win would do so. That probability is much higher. Moreover, this woman has a personal file drawer that contains all those past instances in which she had similar premonitions and then lost.

The same reasoning applies to *precognitive dreams* (in other words, dreams that anticipate an unlikely event that then occurs a few days later). We tend to forget our dreams unless and until an event happens to remind us of them. We thus have no way of evaluating how often we might have dreamed of similar unlikely events that did *not* occur. We fill our database with positive instances and unknowingly exclude the negative instances.

Perhaps the fullest file drawers belong to the so-called psychics who make annual predictions in the tabloid newspapers. Nobody remembers the predictions that fail, but everybody remembers the occasional direct hits. In fact, these psychics are almost always wrong (Frazier, 1987; Tyler, 1977).

Skepticism about Psi

If some of the experimental evidence for psi is as impressive as it seems, why hasn't it become part of established science? Why do we continue to be skeptical?

EXTRAORDINARY CLAIMS Most scientists believe that extraordinary claims require extraordinary proof. A study reporting that students who study harder get higher grades will be believed even if the study was seriously flawed because the data accord well with our understanding of how the world works. But the claim that two people in a ganzfeld study communicate telepathically is more extraordinary; it violates our a priori beliefs about reality. We thus rightly demand a higher measure of proof from parapsychologists because their claims, if true, would require us to radically revise our model of the world—something we should not undertake lightly. In this way, science is justifiably conservative. Many open-minded nonparapsychologists are genuinely impressed by the ganzfeld studies, for example, but reasonably they can and do ask to see more evidence before committing themselves to the reality of psi.

Extraordinariness is a matter of degree. Telepathy seems less extraordinary to most of us than precognition because we are already familiar with the invisible transmission of information through space. We may not all understand how television pictures get to our living rooms, but we know that they do so. Why should telepathy seem that much more mysterious? Precognition, on the other hand, seems more extraordinary because we have no familiar phenomena in which information flows backward in time.

Extraordinariness also depends on our current model of reality. As our understanding of the world changes, a phenomenon that seemed extraordinary at an earlier time may no longer seem so—even if the quality of the evidence has not changed. Any child who has visited a museum of natural history has seen fragments of a meteorite. But before the nineteenth century,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the scientific community did not believe in meteorites. Those who reported seeing them were ridiculed—Stones falling out of the sky? Does God hurl them at us from heaven?—and alternative, natural explanations were advanced to explain away the evidence (Nininger, 1933).

In the twentieth century, quantum mechanics is challenging our everyday model of reality far more radically than most people realize (Herbert, 1987). Some parapsychologists believe that modern physics will provide a model of reality within which psi phenomena will fit comfortably and unremarkably (Stokes, 1987), and many studies of psychokinesis are conducted by physical scientists who explicitly base their theories of psi on quantum mechanics (Jahn & Dunne, 1987). If they are right, the scientific community may come to accept psi not because the data became more convincing but because psi became less extraordinary.

SKEPTICISM OF PSYCHOLOGISTS Psychologists are a particularly skeptical group. National polls find that about one-half of all adult Americans believe in ESP, a figure that rises to two-thirds among Americans with college backgrounds. A survey of over 1,000 college professors found that about 66 percent believe that ESP is either an established fact or a likely possibility. Moreover, these favorable views were expressed by a majority of professors in the natural sciences (55 percent), the social sciences excluding psychology (66 percent), and the arts, humanities, and education (77 percent). The comparable figure for psychologists was 34 percent (Wagner & Monnet, 1979).

Psychologists may be more skeptical than others for several reasons. First, claims of psi might seem more extraordinary to psychologists than to others because it is their conceptual world that would require the most radical revisions if psi were shown to exist. Second, they are the most familiar with past instances of extraordinary claims within psychology that turned out to be based on flawed experimental procedures, faulty inference, or even on fraud and deception.

Third, psychologists know that popular accounts of psychological findings are frequently exaggerated. For example, the genuinely remarkable findings from research on asymmetries in the human brain (see p. 50) have spawned a host of pop-psychology books and media reports containing unsubstantiated claims about left-brained and right-brained persons. Irresponsible reports about states of consciousness—including hypnosis and psi—appear daily in the media. It is thus pertinent to note that when the college professors in the survey cited above were asked to name the sources for their beliefs about ESP, they most frequently cited reports in newspapers and magazines.

And finally, research in cognitive and social psychology has sensitized psychologists to the biases and shortcomings in our abilities to draw valid inferences from our everyday experiences (see Chapter 18). This makes them particularly skeptical of anecdotal reports of psi where, as we saw above, our judgments are subject to many kinds of errors.

For these several reasons, then, much of the skepticism of psychologists toward psi is well-founded. But some of it is not. As we noted earlier, some scientists declare psi to be an impossibility and reject the legitimacy of parapsychological inquiry—a priori judgments that we believe to be out of place in science. Only 4 percent of the college professors in the survey declared psi to be an impossibility—but 34 percent of the psychologists did so. Two hundred years ago, these same skeptics would have been equally certain that God does not hurl stones at us from heaven.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

CHAPTER SUMMARY

1. A person's perceptions, thoughts, and feelings at any moment in time constitute that person's *consciousness*. An *altered state of consciousness* is said to exist when mental functioning seems changed or out of the ordinary to the person experiencing the state. Some altered states of consciousness, such as sleep and dreams, are experienced by everyone; others result from special circumstances, such as meditation, hypnosis, or the use of drugs.
2. The functions of consciousness are a) *monitoring* ourselves and our environment so that we are aware of what is happening within our bodies and in our surroundings; and b) *controlling* our actions so that they coordinate with events in the outside world. Not all events that influence consciousness are at the center of our awareness at a given moment. Memories of personal events and of the knowledge accumulated during a lifetime that are accessible but are not currently part of one's consciousness are called *preconscious memories*. Events that affect behavior even though we are not aware of perceiving them influence us *subconsciously*.
3. According to psychoanalytic theory, some emotionally painful memories and impulses are *not* available to consciousness because they have been repressed—that is, diverted to the *unconscious*. Unconscious thoughts and impulses influence our behavior even though they reach consciousness only in indirect ways through dreams, irrational behavior, and slips of the tongue.
4. The notion of a divided consciousness assumes that thoughts and memories may sometimes be *dissociated*, or split off, from consciousness, rather than repressed to the unconscious. Extreme examples are cases of *multiple personality*, in which two or more well-developed personalities alternate within the same individual.
5. *Sleep*, an altered state of consciousness, is of interest because of the rhythms evident in sleep schedules and in the depth of sleep. These rhythms are studied with the aid of the *electroencephalogram* (EEG). Patterns of brain waves show four stages (depths) of sleep, plus a fifth stage characterized by *rapid eye movements* (REMs). These stages alternate throughout the night. Dreams occur more often during REM sleep than during the other stages (non-REM sleep).
6. In 1900, Sigmund Freud proposed the most influential theory of dreams. It attributes psychological causes to dreams, distinguishing between the *manifest* and *latent content* of dreams and stating that dreams are wishes in disguise.
7. *Psychoactive drugs* have long been used to alter consciousness and mood. They include *depressants*, such as alcohol and tranquilizers; *opiates*, such as heroin and morphine; *stimulants*, such as amphetamines and cocaine; *hallucinogens*, such as LSD and PCP; and *cannabis*, such as marijuana and hashish.
8. All of these drugs can produce *psychological dependence* (compulsive use to reduce anxiety), and most result in *physical dependence* (increased tolerance and withdrawal symptoms) if used habitually.
9. Alcohol is an integral part of social life for many college students, but it can create serious social, psychological, and medical problems. Prolonged heavy drinking can lead to *alcoholism*, which is marked by an *inability to abstain* from or a *lack of control* over drinking.
10. *Meditation* represents an effort to alter consciousness by following planned rituals or exercises such as those of yoga or Zen. The result is a somewhat mystical state in which the individual is extremely relaxed and feels divorced from the outside world. Simple exercises combining concentration and relaxation can help novices experience meditative states.
11. *Hypnosis* is a responsive state in which subjects focus their attention on the hypnotist and the hypnotist's suggestions. Some people are more readily hypnotized than others, though most people show some susceptibility. Self-hypnosis can be learned by those who are responsive to hypnosis induced by others.
12. Characteristic hypnotic responses include enhanced or diminished *control over movements*, the distortion of memory through *posthypnotic amnesia*, *age*

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

regression, and positive and negative *hallucinations*. The reduction of pain, as a variety of negative hallucination, is one of the beneficial uses of hypnosis in the treatment of burns and in obstetrics, dentistry, and surgery.

13. Theories of hypnosis have long been a source of controversy, with each explaining some aspect of hypnotic behavior but none explaining all. With better agreement on the empirical facts, the theories are gradually becoming complementary rather than antagonistic.
14. There is considerable controversy over *psi*, the idea that human beings can acquire information about the world in ways that do not involve stimulation of known sense organs or can influence physical events by purely mental means. The phenomena of psi includes *extrasensory perception* (ESP) in its various forms (telepathy, clairvoyance, precognition) and *psychokinesis*, movement of objects by the mind.
15. A number of carefully controlled studies (called *ganzfeld experiments*) have been conducted to evaluate ESP via telepathy. These experiments are subject to criticism (replicability, inadequate controls, file-drawer problems). However, a careful analysis of the results does not preclude the possibility of a real ESP effect. Nevertheless, most psychologists remain skeptical about ESP and psi in general, in part because so many past instances of extraordinary claims turned out to be based on flawed experimental procedures, faulty inferences, or even on fraud and deception.

FURTHER READING

Several books deal in general with the problems of consciousness and its alterations, such as Baars, *Cognitive Theory of Consciousness* (1988); Pope and Singer (eds.), *The Stream of Consciousness* (1978); and Bowers and Meichenbaum (eds.), *The Unconscious Reconsidered* (1984). For philosophical/psychological discussions of consciousness see Lycan, *Consciousness* (1987); Jackendoff, *Consciousness and the Computational Mind* (1987); and Churchland, *Matter and Consciousness* (1988).

Problems of divided consciousness are treated in Hilgard, *Divided Consciousness* (1977); Kluft (ed.), *Childhood Antecedents of Multiple Personality* (1985); and Braun (ed.), *Treatment of Multiple Personality Disorder* (1986).

Useful books on sleep and dreams include Hobson, *Sleep* (1989); Horne, *Why We Sleep* (1988); Hobson, *The Dreaming Brain* (1988); and Hauri, *The Sleep Disorders* (1982).

General textbooks on drugs include Julien, *A Primer of Drug Action* (5th ed., 1988); Julien, *Drugs and the Body* (1988); and Ray, *Drugs, Society, and Human Behavior* (3rd ed., 1983). *Drug and Alcohol Abuse* (3rd ed., 1989) by Schuckit provides a guide to diagnosis and treatment. For a thoughtful discussion of the legal and social problems of heroin, as well as an evaluation of possible solutions, see Kaplan, *The Hardest Drug; Heroin and Public Policy* (1983).

On meditative practices, see West (ed.), *The Psychology of Meditation* (1987); Goleman, *The Varieties of Meditative Experience* (1977); or Naranjo and Ornstein, *On the Psychology of Meditation* (1977). On meditation for relaxing, see Benson, *The Relaxation Response* (1975). For a discussion of relaxation and mental images in athletics see Syer and Connolly, *Sporting Body Sporting Mind; An Athlete's Guide to Mental Training* (1984) and Butt, *The Psychology of Sport* (2nd ed., 1987).

There are a number of books on hypnosis. Presentations that include methods, theories, and experimental results are E. R. Hilgard, *The Experience of Hypnosis* (1968); Fromm and Shor (eds.), *Hypnosis; Developments in Research and New Perspectives* (2nd ed., 1979); and J. R. Hilgard, *Personality and Hypnosis* (2nd ed., 1979).

For a review of parapsychology, see Wolman, Dale, Schmeidler, and Ullman (eds.), *Handbook of Parapsychology* (1985); Frazier (ed.), *Science Confronts the Paranormal* (1986); Kurtz (ed.), *A Skeptic's Handbook of Parapsychology* (1985); Marks and Kammann, *The Psychology of the Psychic* (1980); and Gardner, *Science: Good, Bad, and Bogus* (1981).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Physica B 151 (1988) 339–348
North-Holland, Amsterdam

TESTING SCHRÖDINGER'S PARADOX WITH A MICHELSON INTERFEROMETER

Evan Harris WALKER
U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, USA

E.C. MAY, S.J.P. SPOTTISWOODE and T. PIANTANIDA
SRI International, Menlo Park, California, USA

The Schrödinger paradox points out that quantum mechanics predicts a linear superposition of states even for macroscopic objects prior to measurement. However, at the macroscopic level of ordinary objects it has not been possible to maintain the phase correlations needed to demonstrate or disprove the reality of such a superposition of states as opposed to the mixture of states. Without such a quantum "signature", this paradoxical prediction of quantum theory would seem to have no testable consequences. State vector collapse in that case becomes indistinguishable from a stochastic ensemble description.

The experiment described here provides a means for testing Schrödingers' paradox. A Michelson interferometer is used to test for the presence of state superposition of a pair of shutters that are placed along the two optical arms of the interferometer and driven by a beta decay source so that either the first shutter is open and the second closed or *vice versa*. The shutters take on the role of the cat in the Schrödinger paradox.

The experiment that we discuss here has been carried out at SRI International. Under the conditions of the experiment, the results remove the possibility of the existence of macroscopic superposition prior to observation.

1. Introduction

The Schrödinger paradox is among the oldest of the puzzles surrounding the interpretation of quantum mechanics. Like the Einstein–Podolsky–Rosen (EPR) paradox it has engendered a great deal of speculation about our basic understanding of physical reality. Also like the EPR paradox, Schrödinger proposed his paradox to point out that the statistical interpretation of quantum theory must at some level contain a flaw, since it implies the reality of a linear superposition of states even at the macroscopic level – before observation. Moreover, just as in the case of the EPR paradox, the Schrödinger paradox has long been thought of as an untestable consequence of quantum theory, since it relates to the state of a macrosystem just before observation, a state that we *know* must approach asymptotically to that given by classical mechanics. That is to say, we know that the usual interference effects by which we distinguish the presence of state superposition in atomic processes can be shown to be too small to observe in the case of macroscopic systems. But the development of Bell's theorem showed us how to test the paradoxical implications of quantum mechanics that had been pointed out by Einstein, Podolsky and Rosen in 1935. Within the limits of our experimental setup, we have now done the same for the Schrödinger paradox.

There are important reasons for doing this experiment. All of us are quite aware of the fact that the existence of a linear superposition of states at the macroscopic level is quite counterintuitive. Nevertheless, no experiment has ever been done that has yielded results contrary to the literal application of quantum theory. The absence of superposition at the macrolevel *prior to observation* has not been experimentally demonstrated – and in fact it has generally been thought that such a test was not feasible. This has led to the developement of various interpretations of quantum mechanics having to do with the macroscopic reality of quantum states.

A second reason for carrying out an experiment of the present type is that it represents an efficient way to search for the nature of and

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

cause of state vector collapse. We all know that the machinery that effects state vector collapse, whatever that phrase actually means, must be somewhere between the thing observed and the observer. Indeed, this observer–observed dichotomy has become a rather commonly used phrase in discussions of the measurement problem. But the gap between these two covers a lot of territory. Moreover, we also know that the perponderant opinion is that the transition from the pure state to the mixed state probably takes place at a level above that of the largest coherence that exists for the system being observed. But to focus all our attention at that level at this stage of the game when we still know so little about what causes state vector collapse may not be an efficient way to explore the physics involved. It may be a more efficient strategy to carry out experiments looking for the existence of state superposition at various levels between the atomic level and that of the macroscopic world. Most experiments in this field are designed to examine a cut in von Neumann's chain between the observer and the observed just above the level of the basic atomic interaction itself. Our experiment goes to the opposite extreme to look for state superposition immediately prior to observation at the macroscopic level.

By doing this we are able to deal experimentally with what has come to be a quite widespread and popular conception of what quantum mechanics has to say about physical reality. The Schrödinger paradox has been used to imply an actual "observer–observed" dichotomy exists as a fundamental aspect of physical reality, and to imply that the observer creates his own reality in the act of observation. It has been used to raise such questions as the "Wigner's friend" paradox and even to promote speculation that by our observation we may be creating the Big Bang of the universe. If our experiment does nothing more than lay such speculation to rest it will have been more than worthwhile.

At the other extreme, however, we should recognize the possibility that it is through this doorway that some phenomena, heretofore not dealt with by science, may be approached. The existence of consciousness as a phenomenology that lies beyond what we as physicists mean by distance, mass, electric charge and the other constructs of our physical equations cannot be denied. Consciousness surely arises out of something that goes on in the brain of each of us, but yet lies beyond its usual description as a physical object no matter how complex. It may be that if quantum mechanics does require the *observer* as an essential and irreducible aspect of physical reality, then we may find its proper scientific description to be bound up with an understanding of how state vector collapse comes about. Whatever the likelihood that we will find evidence for this in the experiment we discuss here, it would seem to be worth the effort to look.

The Schrödinger paradox arises because the prescription for writing the general state vector for any system requires that one can sum all possible component states for an unmeasured or unobserved system irrespective of the scale of the system to be observed. As a consequence according to Schrödinger, a cat placed in a box rigged to release a tranquilizer (out of difference to the SPCA – and this writer) if a beta decay occurs in a specified interval of time, *or not* if the beta decay does not occur, must be represented by a state vector that is the sum of both possible outcomes before a measurement is made on the system. Using Dirac's notation this would give for the combined state

$$|\Psi\rangle = |\psi_A\rangle + |\psi_S\rangle\ , \qquad (1)$$

where the subscripts A and S refer to the awake and sleep states respectively. Although this is generally regarded to be a preposterous conclusion, there exist no experiments that violate this or any of the basic premises of quantum mechanics. A definitive experiment that would demonstrate that such a superposition of states does not exist would be a significant if not a surprising achievement. On the other hand, since there exist no examples of a violation of the principles of quantum mechanics, it must be considered a viable possibility that quantum mechanics is valid here as well.

It is usually thought that one of two possible

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

influences causes state vector collapse. One of these is that something happens during the transition from the microscopic realm of the system to the macroscopic realm. Efforts so far to formulate such a suggestion have been unsuccessful – in fact all such proposals that have been reduced to a mathematical prescription have proved to be wrong because they have predicted results at variances with experiments already conducted. Of course we know that on measurement state vector collapse will occur – or will have occurred. If we open the box, we will see the cat either awake or asleep. Some few scientists have suggested that the act of conscious observation causes state vector collapse. Wigner has pointed out that such is a peculiar implication of quantum mechanics, but he has made no effort to formulate what this would mean, if indeed he takes this possibility seriously. Wheeler has pointed out that Bohr specifically "rejected the term 'consciousness' in describing the elemental act of observation . . . he emphasized that no measurement is a measurement until it is 'brought to a close by an irreversible act of amplification and

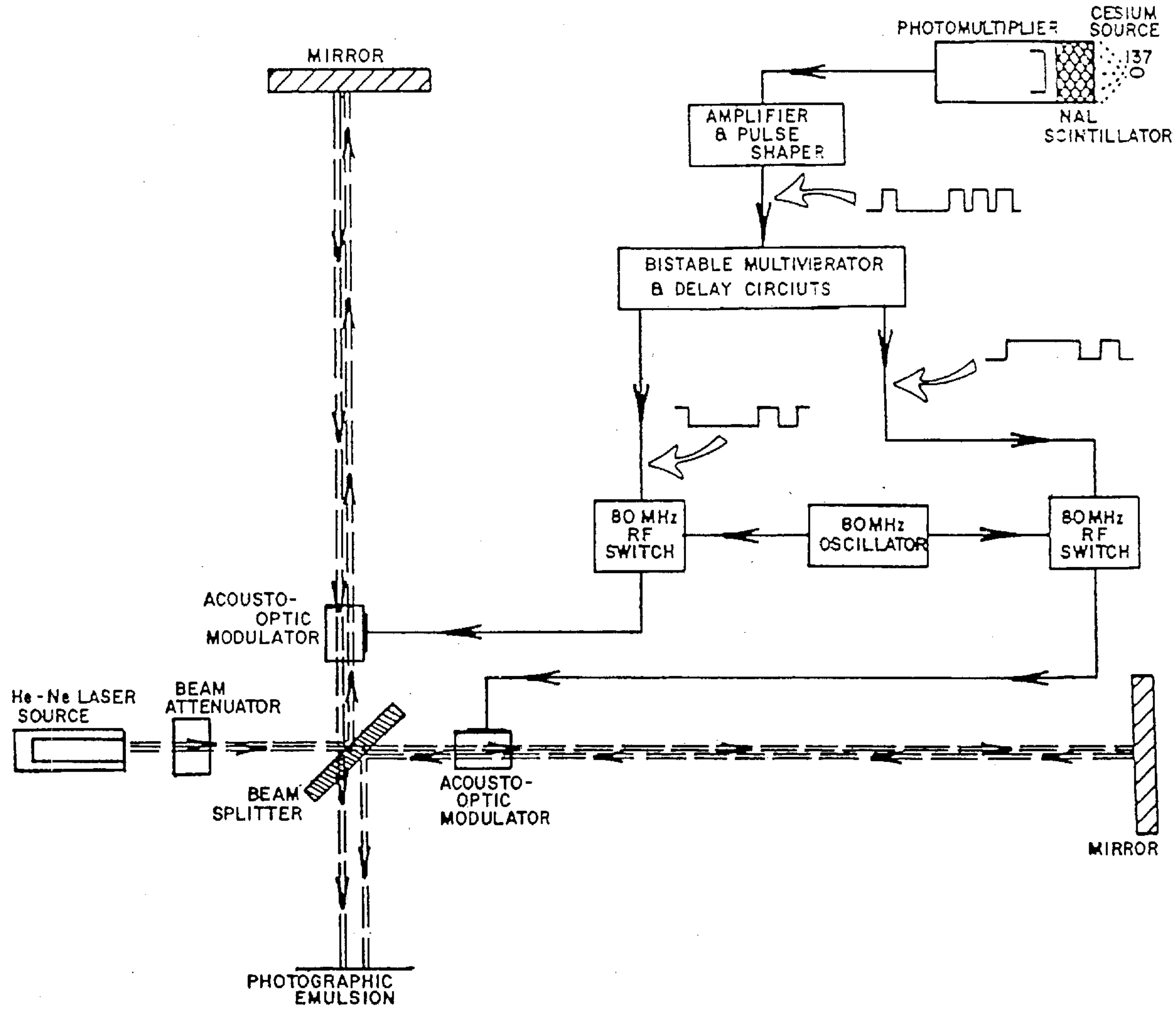


Fig. 1. Experimental arrangement for testing the Schrödinger paradox using a Michelson interferometer. The cesium 137 gamma source provides the random quantum event that triggers the bistable multivibrator (flip-flop) circuit controlling the AO cells so that in any given state one cell is always on while the other is always off.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

until the result is 'communicable in plain language.'" These ideas have not been reduced to mathematical formulation, have not been derived from the Schrödinger equation (note that the Ehrenfest theorem does not cover the case considered here; it does not show that any reduction in the number of states exists in the transition to the microscopic, only that certain kinds of macroscopic processes approach the classical in the limit), and the latter prescription is clearly anthropomorphic – nothing more.

We have carried out an experiment that not only tests the Schrödinger paradox, but has the potential through modest modifications to range the entire gamut of possibilities in order to establish exactly where and how state vector collapse takes place. The experiment makes use of Michelson interferometer (a Mach–Zehnder interferometer could equally be used) in which two shutters, acousto-optical (AO) cells, are placed one in each arm of the interferometer. These AO cells are driven by a quantum mechanical process, specifically, a cesium 137 gamma source driving a bistable multivibrator (flip–flop) circuit in such a way as to gate one or the other of the two possible paths in the interferometer. Thus, knowing the state of the quantum process driving the AO cells we would know that either path 1 was open while path 2 was closed or vice versa. Since we do not know the quantum state driving the AO cells, however, the system must be in both states – that is, in the linear superposition of states. Fig. 1 shows the layout of the experiment.

Although we have replaced Schrödinger's cat with the more manageable AO cells, it is easy to see – as in figs. 2 and 3 – that this is a realization of the Schrödinger paradox in which we have

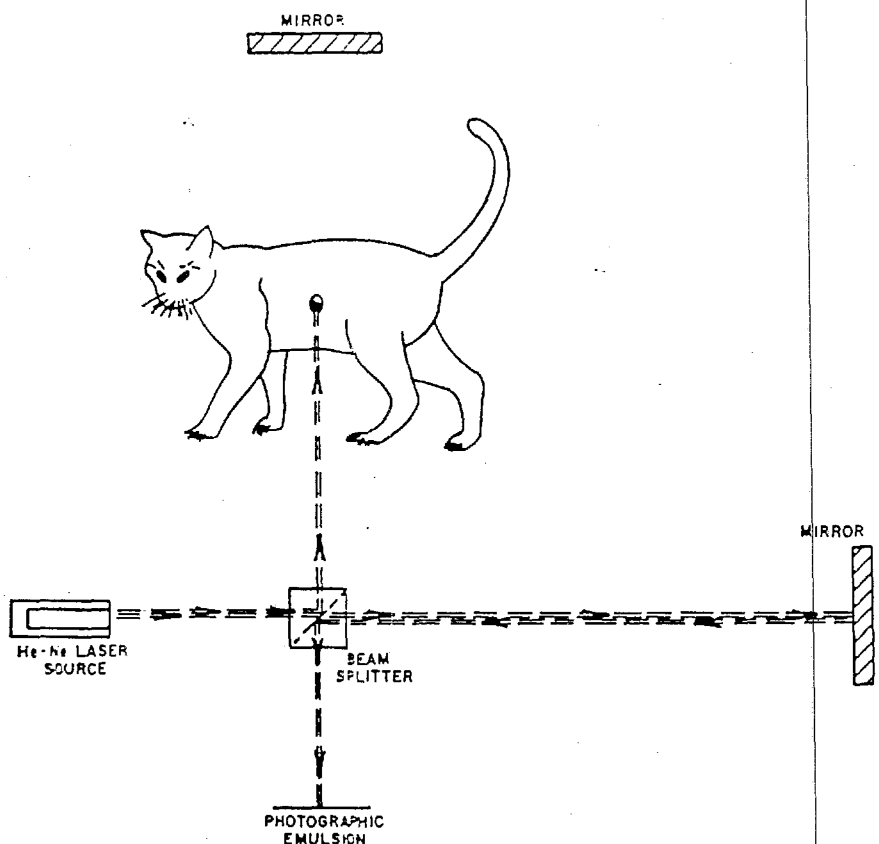


Fig. 2. Schematic showing that the experiment of fig. 1 is in fact a variant of the Schrödinger paradox arrangement. Here the cat in Schrödinger's thought experiment is shown awake and sitting in the way of arm 1 of the interferometer.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

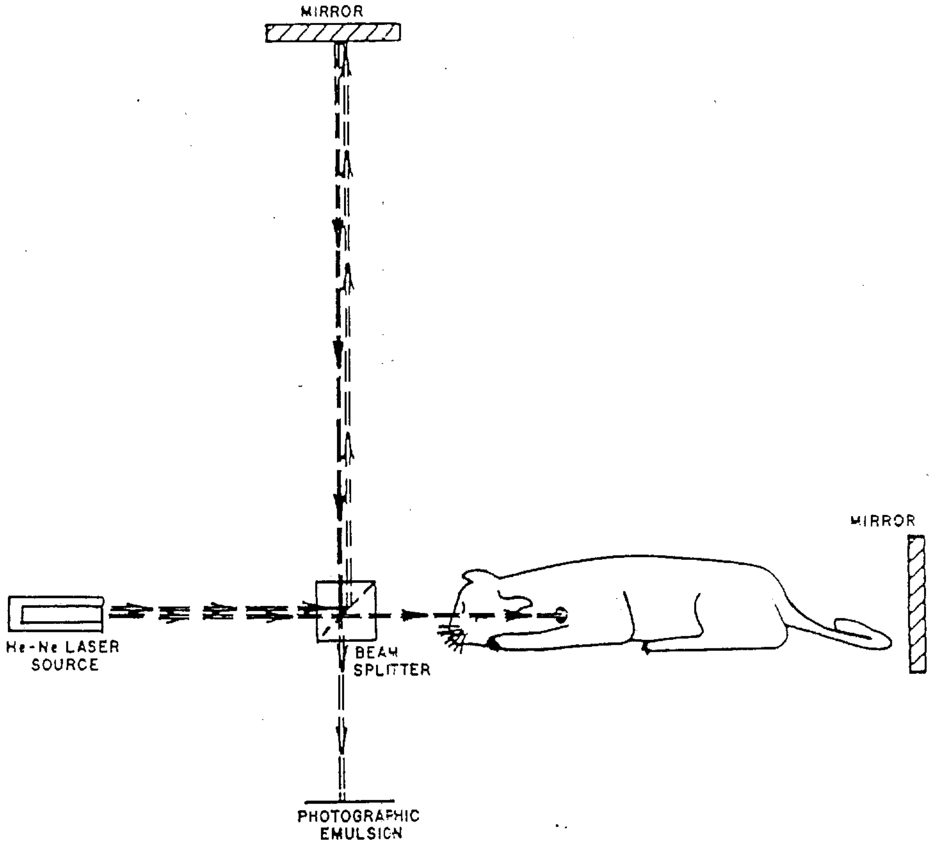


Fig. 3. In this figure we see the state in which the beta source has caused the release of the tranquilizer gas – causing the cat to fall asleep in the way of arm 2 of the interferometer.

added an interferometer to test the existence of a superposition of pure states or simply the presence of one or the other unknown state of a mixture of states. Now let us look at the equations appropriate to the problem.

2. The cat and the correlation of states

Let us now look at why we should not ordinarily expect to observe any effect with large objects in the first place. The system described by $|\Psi\rangle = |\psi_1\rangle + |\psi_2\rangle$ for which an observation operator $A(x)$ in configuration space would yield $A|\psi_1\rangle = \beta_1|\psi_1\rangle$ and $A|\psi_2\rangle = \beta_2|\psi_2\rangle$ yields for an observation probability p,

$$\begin{aligned} p &= \langle\Psi|A^{\dagger}A|\Psi\rangle \\ &= (\beta_1^*\langle\psi_1| + \beta_2^*\langle\psi_2|)(\beta_1|\psi_1\rangle + \beta_2|\psi_2\rangle) \\ &= |\beta_1|^2 + |\beta_2|^2 + \beta_1^*\beta_2\langle\psi_1|\psi_2\rangle \\ &\quad + \beta_2^*\beta_1\langle\psi_2|\psi_1\rangle \,. \end{aligned} \tag{2}$$

Because our object is large however, the phase factors entering into ψ_1 and ψ_2 will vary rapidly, so rapidly that the terms $\langle\psi_1|\psi_2\rangle$ and $\langle\psi_2|\psi_1\rangle$ are for macroscopic objects zero. As a consequence, Eq. (2) reduces to $p = |\beta_1|^2 + |\beta_2|^2$ which is indistinguishable from the classical probabilities for the system.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

3. Michelson interferometer

Assume the arrangement shown in fig. 1, but in which both AO cells are always on. This gives us then the standard Michelson interferometer with a CW laser source. Using the subscripts 1 and 2 for the two arms of the device, we write for the state of a single photon $|\Psi_0\rangle = |\phi_1\rangle + |\phi_2\rangle$. For a configuration space photon absorption operator $A(x)$ satisfying $A(x)|\phi_1\rangle = \beta_1|\phi_1\rangle$ and $A(x)|\phi_2\rangle = \beta_2|\phi_2\rangle$ obviously we will have for the probability p_0, $p_0 = \langle\Psi_0|A^{\dagger}A|\Psi_0\rangle$, so that

$$p_0 = |\beta_1|^2 + |\beta_2|^2 + \beta_1^*\beta_2\langle\phi_1|\phi_2\rangle + \beta_2^*\beta_1\langle\phi_2|\phi_1\rangle\,, \tag{3}$$

which is formally what we found in eq. (2). However, for photons in an interferometer, terms like $\langle\phi_1|\phi_2\rangle$ contribute significantly. We will return to this later. The point here, however, is that it is the presence of these cross terms that lead to the interference effects we observe with an interferometer.

Since we are using a constant wave (CW) laser for our source, the photon is represented by

$$\phi_1 = a_1\,e^{i(kx_1-\omega t_1)}\,, \tag{4}$$

where x_1 is the path length for arm number one in the interferometer, t_1 is the time of the measurement, k and ω are the wave number and angular frequency and a_1 is a normalization factor. If $\langle\phi_1|\phi_2\rangle$ and $\langle\phi_2|\phi_1\rangle$ are averaged over complete cycles of x_1, x_2, t_1 and t_2, these terms will vanish. With equivalent paths for the two arms of the interferometer, $|\beta_1|^2 = |\beta_2|^2 = \beta^2$ so that we can simply define $\bar{p}_0 = 2\beta^2$.

4. Michelson interferometer test of the Schrödinger paradox

Now let us look at the the complete problem as shown in fig. 1 in which the state of the quantum mechanical system depends on the coupled gamma decay–photon system. The gates, G, are functions of a parameter B of the gamma decay and of the arm of the interferometer i[n] which the gate is located, while the photon re[-]presentation as before depends on the arm of [the] interferometer, position and time t. We have [in] general $|\Psi\rangle = |G\rangle\otimes|\Phi\rangle$. This gives us

$$|\Psi\rangle = G_1(B,\mathrm{arm}\,1)|\phi_1(x_1,t)\rangle + G_2(B,\mathrm{arm}\,2)|\phi_2(x_2,t)\rangle\,. \tag{5}$$

The parameter B has two states which we desig[-]nate "on" (or "+") and "off" (or "−") for arm number one of the interferometer and for th[e] second arm, "off" (or "−") and "on" (or "+") respectively, as determined by the logic of [the] switching circuit. Eq. (5) becomes

$$|\Psi_1\rangle = G_1^+|\phi_1\rangle + G_2^-|\phi_2\rangle + G_1^-|\phi_1\rangle + G_2^+|\phi_2\rangle\,. \tag{6}$$

As before we write $A(x)|\phi_1\rangle = \beta_1|\phi_1\rangle$. etc. [We] therefore have

$$A|\Psi_1\rangle = G_1^+\beta_1|\phi_1\rangle + G_2^-\beta_2|\phi_2\rangle + G_1^-\beta_1|\phi_1\rangle + G_2^+\beta_2|\phi_2\rangle\,. \tag{7}$$

The detection probability function p_1 is then

$$\begin{aligned} p_1 &= \langle\Psi_1|A^{\dagger}A|\Psi_1\rangle \\ &= \langle(G_1^+\beta_1\phi_1 + G_2^-\beta_2\phi_2 + G_1^-\beta_1\phi_1 + G_2^+\beta_2\phi_2)|(G_1^+\beta_1\phi_1 + G_2^-\beta_2\phi_2 + G_1^-\beta_1\phi_1 + G_2^+\beta_2\phi_2)\rangle\,, \end{aligned} \tag{8}$$

which gives

$$\begin{aligned} p_1 = {} & |\beta_1|^2|G_1^+|^2 + \beta_1^*\beta_2 G_1^{+*}G_2^-\langle\phi_1|\phi_2\rangle \\ & + |\beta_1|^2 G_1^{+*}G_1^- + \beta_1^*\beta_2 G_1^{+*}G_2^+\langle\phi_1|\phi_2\rangle \\ & + \beta_2^*\beta_1 G_2^{-*}G_1^+\langle\phi_2|\phi_1\rangle + |\beta_2|^2|G_2^-|^2 \\ & + \beta_2^*\beta_1 G_2^{-*}G_1^-\langle\phi_2|\phi_1\rangle + |\beta_2|^2 G_2^{-*}G_2^+ \\ & + |\beta_1|^2 G_1^{-*}G_1^+ + \beta_1^*\beta_2 G_1^{-*}G_2^-\langle\phi_1|\phi_2\rangle \\ & + |\beta_1|^2|G_1^-|^2 + \beta_1^*\beta_2 G_1^{-*}G_2^+\langle\phi_1|\phi_2\rangle \\ & + \beta_2^*\beta_1 G_2^{+*}G_1^+\langle\phi_2|\phi_1\rangle + |\beta_2|^2 G_2^{+*}G_2^- \\ & + \beta_2^*\beta_1 G_2^{+*}G_1^-\langle\phi_2|\phi_1\rangle + |\beta_2|^2|G_2^+|^2\,. \end{aligned} \tag{9}$$

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Since the component states such as ϕ_1 and ϕ_2 in the linear superposition are the same functions as for the single states alone, we have simply that $G_1^- = G_2^- = G_1^{-*} = G_2^{-*} = 0$, and $|G_1^+|^2 = |G_2^-|^2 = 1$. Eq. (9) reduces to

$$p_1 = |\beta_1|^2 + |\beta_2|^2 + \beta_1^* \beta_2 \langle \phi_1 | \phi_2 \rangle + \beta_2^* \beta_1 \langle \phi_2 | \phi_1 \rangle . \quad (10)$$

which is the same result as in eq. (3) for the Michelson (or Mach–Zehnder) interferometer result without AO cells.

5. Representation of the photon

Since the laser is not pulsed and since we can assume the switching rate of the AO cells to be much lower than the frequency of the photon, we can write simply

$$A\phi_1 = \beta_1 a_1 \, e^{i(kx_1 - \omega t)} \quad (11)$$

and

$$A\phi_2 = \beta_2 a_2 \, e^{i(kx_2 - \omega t)} , \quad (12)$$

where we have introduced the factors a_1 and a_2 as normalization factors for the particular conditions of the experimental arrangement and detection interval. For essentially identical arms in the interferometer, $\beta_1 a_1 = \beta_2 a_2$, so that we can write

$$p_0' = \Psi | A^\dagger A | \Psi \rangle' = \alpha^2 \Big[2\Delta x + \int_{x_1}^{x_1+\Delta x} e^{ik(x_2 - x_1')} \, dx_1' + \int_{x_2}^{x_2+\Delta x} e^{ik(x_1 - x_2')} \, dx_2' \Big] , \quad (13)$$

where Δx is the thickness of the photographic film layer and where we have incorporated the time interval of the measurement in our normalization factor α. The primes indicate that the probabilities represent a measurement over a time that is long with respect to the photon frequency. Of course, for thin films we have simply

$$p_0' = \alpha^2 [2 + e^{ik(x_2 - x_1)} + e^{ik(x_1 - x_2)}] . \quad (14)$$

Therefore, in the absence of any formalism that would prescribe state vector collapse below the macroscopic level, our calculation predicts the presence of interference fringes in the present experiment despite its counterintuitiveness.

Therefore, a failure to detect a robust interference pattern will show that the experimental results are in disagreement with our theoretical prediction.

The converse outcome holds equally remarkable significance. The occurrence of interference fringes would mean that the linear superposition of state holds *before observation* even on the macroscopic scale.

6. The apparatus

The apparatus consists of a simple Michelson interferometer with optical switches in the relay arms. A schematic diagram of the arrangement is shown in fig. 1.

The polarized output beam from a 6328 Å CW helium–neon single mode laser is attenuated by a factor of 10^{-9} so as to produce a beam of 1.3×10^{-14} W, approximately 4.17×10^4 photons per second intensity. The attenuation is achieved by a combination of the deflected beam intensity reduction, neutral density filters and a polarizer. The light incident on the beam splitting is polarized with the electric vector perpendicular to the plane of fig. 1.

The light is passed to a beam splitter which produces beams of nearly equal intensity. Each arm of the interferometer contains an acousto-optical modulator consisting of a TeO_2 crystal coupled to an ultrasonic piezoelectric transducer. When no input voltage is applied to the transducer, light travels through the crystal undeviated. With an 80 MHz rf signal applied to the transducer, the resulting acoustic waves in the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

crystal diffract the light beam by twice the Bragg angle, which is 5.9 mrad for the devices used. It is this diffracted beam that is used to obtain interference in the interferometer. Turning the transducer off turns the AO cell off. These acousto-optical devices exhibit a rise and fall time of 120 ns when switching a beam of 0.75 mm diameter. Most of the power (80%) goes into the first-order diffracted beam, which leaves the device at an angle of 11.8 mrad to the zero-order undiffracted beam, while smaller amounts of power are deposited into the higher orders. As stated, it is the first-order diffracted beams that are reflected off the interferometer mirrors and back through the modulators where a second diffraction occurs. Only the first-order beams are shown in fig. 1 for clarity. The zero-order and higher-order diffacted light is absorbed by various beam stops. Thus the acousto-optical modulators act to chop the light in the interferometer arms with a contrast ratio adequate to assure that none of the photons reaching the film will have passed through an "off" shutter.

The switched beams from the interferometer arms are recombinated in the beam splitter and are deposited on a high speed photographic emulsion. The beam divergence of the laser and the geometry of the apparatus are chosen so as to result in a 2 mm diameter image on the film with approximately four linear fringes visible in this area.

The acousto-optical modulators are driven by switched 80 MHz oscillators which are in turn gated on and off by the outputs of a bistable multivibrator, or flip–flop, so that one modulator is on while the other is off. A delay of 150 ns is introduced into the gating signals applied to the rf drivers so that one shutter does not start to open until the other has closed, thus ensuring that at no time are both interferometer arms open. The flip–flop is clocked by pulses from a photomultiplier, which have been suitably amplified, shaped and discriminated. The photomultiplier looks into a sodium iodide scintillator crystal. With a cesium 137 source of approximately 30 μCi placed 2 cm from the scintillator, the pulses which clock the flip–flop have a mean repetition rate of 118 kHz.

With an average photon rate of $4.17 \times 10^4\ s^{-1}$ emerging from the attenuator, there is an average of less than one photon in each arm of the

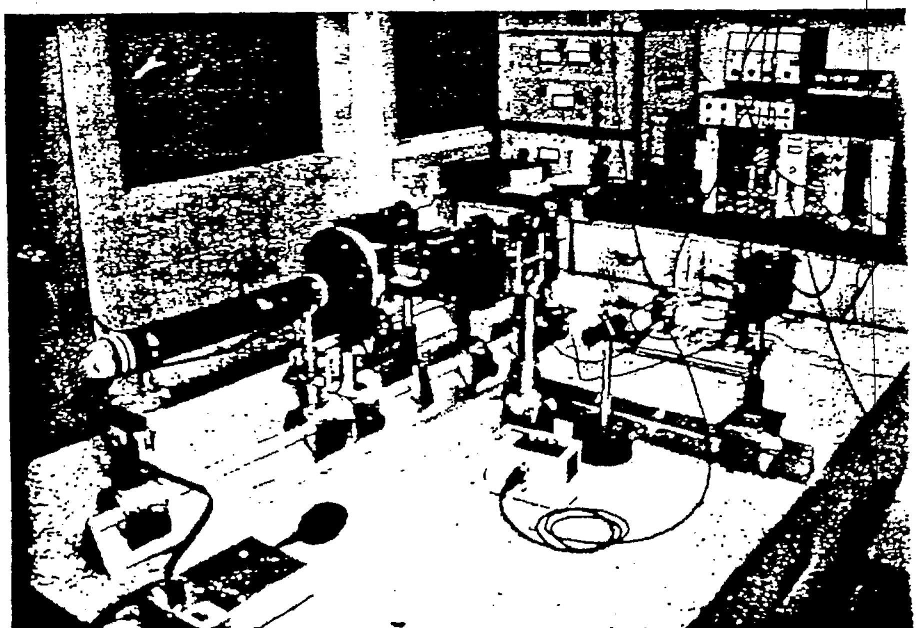


Fig. 4. Photograph of the laboratory layout.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

interferometer during each period for which the acousto-optic modulator in the arm is open. The resulting low intensity image is recorded on Kodak 2415 technical pan film hypersensitised in forming gas prior to exposure. A high speed Polaroid film has also been used. A photograph of the laboratory setup is shown in fig. 4.

7. Test runs

In order to verify that the apparatus properly discriminates between the two possible experimental outcomes, test runs were made with the logic circuit connected so that both shutters would either be on or off to make certain that the apparatus could give interference patterns. This run gives us a reference interference pattern that we can use to judge the results of our experimental run. The average shutter rate in this case was the same as for the final experimental run. The resulting photograph is shown in Fig. 5.

8. Experimental results and conclusions

Fig. 6 shows the experimental outcome for a photon rate of $4.17 \times 10^4\ s^{-1}$. The figure speaks for itself. Thee is no interference pattern present in the figure. The figure clearly shows the absence of the interference fringes predicted by our formalism.

This absence of interference fringes in the experimental run is a result that, although expected on the basis of commonsense, we nevertheless interpret as a *prima facie* case that quantum theory may be violated. The result is important because it provides a starting point for us in our search for the cause and experimental meaning of state vector collapse. These results are also important because they are related to questions about the Schrödinger paradox.

We do not yet know just how and where state vector collapse occurs – nor do we know what this even means. Our experiment does not solve or remove the measurement problem. If anything, it deepens the problem. We must find out

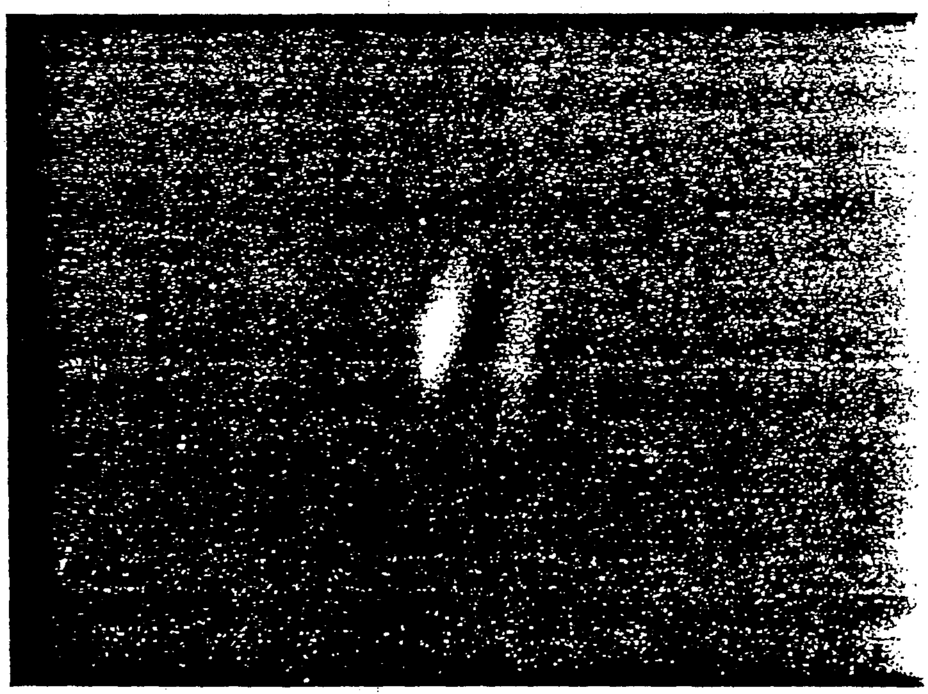


Fig. 5. Test run in which both shutters are either open or closed at the same time to assure a conventional interference pattern. Photon rate was $4.17 \times 10^4\ s^{-1}$, less than one photon in each arm of the apparatus at any moment.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

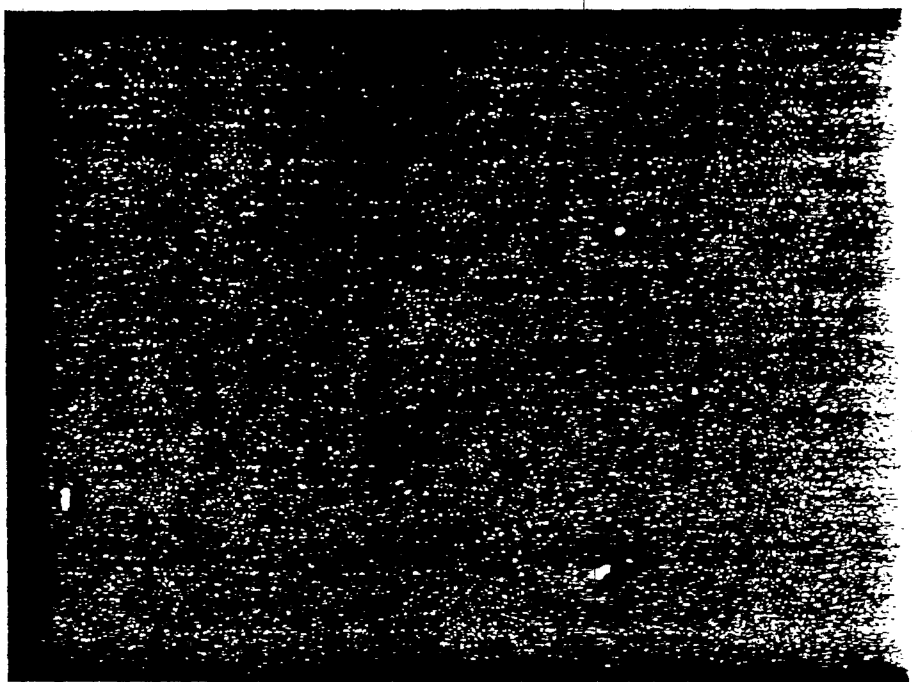


Fig. 6. Final experimental run. No interference pattern was obtained. This figure clearly shows the absence of interference.

the mechanics of state vector collapse. The present experiment provides a basic plan for future experiments to search out how state vector collapse occurs and to give us a clear understanding of just what state vector collapse entails. It gives us the tool we need to find just where to cut von Neumann's chain.

Acknowledgements

We wish to acknowledge the helpful discussions about this experiment held with C.O. Alley, Y.H. Shih, and George Hinds of the Department of Physics of the University of Maryland.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Parapsychological Research: A Tutorial Review and Critical Appraisal

RAY HYMAN

Invited Paper

Beginning in the 1850s, some eminent scientists such as Robert Hare, Alfred Russel Wallace, and Sir William Crookes investigated the claims of spiritualist mediums and believed that they had demonstrated scientifically the existence of psychic phenomena. Critics, without examining the evidence, dismissed the claims out of hand and charged the offending scientists with gross incompetence or with fraud. Encouraged by the work of these early psychical researchers, a group of scholars founded the Society for Psychical Research in London in 1882. In spite of this beginning, psychical research remained an amateur and uncoordinated set of activities until the publication of Rhine's Extra-Sensory Perception *in 1934. The card-guessing experiments featured in Rhine's book became the model for experimental parapsychology for the next 40 years. Since the 1970s Rhine's paradigm has been replaced by a number of research programs such as remote viewing, the Ganzfeld experiment, and psychokinetic investigations using Random Event Generators. The present paper examines examples of what were considered, in their time, the best examples of scientific evidence for paranormal phenomena. Each generation of parapsychologists has set aside the work of earlier generations and offered up as sufficient scientific evidence the best work of its own day. As a result, parapsychology lacks not only lawful and replicable phenomena, but also a tradition of cumulative evidence. Two systematic evaluations of the best contemporary research programs in parapsychology revealed that the experiments departed from the minimal standards of adequate randomization of targets, appropriate use of statistical inference, and controls against sensory leakage. The historical survey in this paper suggests that the same themes and inadequacies that haunted the very earliest investigations still characterize contemporary parapsychological research. Both proponents and critics throughout the 130 years of the controversy over psychical research, have deviated greatly from those standards of fair-play and rationality that we would like to believe characterizes the best scientific arguments. Some encouraging signs for progress towards resolving some of the issues raised by the controversy have recently appeared. The criticism of the parapsychological claims is becoming more informed and constructive. Many younger parapsychologists have been working for higher standards within their field. The best lines of systematic research in parapsychology are not of sufficient quality to be put before the scrutiny of the rest of the scientific community. However, with the recent increase in constructive criticism and with the growing awareness within the parapsychological community that it needs to specify minimal standards and set its own house in order, there is hope that in the near future either the parapsychologists will fail to find evidence for psi or will be ready to challenge the scientific community with the sort of evidence that it cannot ignore.*

Manuscript received January 25, 1985; revised August 21, 1985.
The author is with the Psychology Department, University of Oregon, Eugene, OR 97403, USA.

INTRODUCTION

Robert Jahn, Dean of the School of Engineering and Applied Science at Princeton University, can be taken as a representative example of what happens when an eminent and established scientist takes the time to carefully examine the evidence for paranormal phenomena. About seven years ago, an undergraduate requested him to supervise her investigation of psychic phenomena [1].

> Although I had no previous experience, professional or personal, with this subject, for a variety of pedagogical reasons I agreed, and together we mapped a tentative scholarly path, involving a literature search, visits to appropriate laboratories and professional meetings, and the design construction, and operation of simple experiments. My initial oversight role in this project led to a degree of personal involvement with it, and that to a growing intellectual bemusement, to the extent that by the time this student graduated. I was persuaded that this was a legitimate field for a high technologist to study and that I would enjoy doing so.

As a result of his own survey of the field as well as his own initial experiments in parapsychology, Jahn concludes that [1]:

> once the illegitimate research and invalid criticism have been set aside, the remaining accumulated evidence of psychic phenomena comprises an array of experimental observations, obtained under reasonable protocols in a variety of scholarly disciplines, which compound to a philosophical dilemma. On the one hand, effects inexplicable in terms of established scientific theory, yet having numerous common characteristics, are frequently and widely observed; on the other hand, these effects have so far proven qualitatively and quantitatively irreplicable, in the strict scientific sense. and appear to be sensitive to a variety of psychological and environmental factors that are difficult to specify, let alone control.

Jahn, like many of his predecessors who took a serious look at the evidence for the paranormal, finds the phenomena to be erratic, evasive, and ephemeral. Indeed, he admits that when judged according to strict scientific standards the evidence for the actual existence of the phenomena is not "fully persuasive." But he is intrigued. Like his predecessors, he is optimistic that with the right application of technology and scientific ingenuity the phenomena can be captured and made lawful.

0018-9219/86/0600-0823$01.00 ©1986 IEEE

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

This is one of a number of justifiable reactions one can have as a result of fairly examining the case for psychical research. Jahn is willing to risk his time and reputation on the possibility that careful and diligent investigation will bring some lawfulness to this unruly area of inquiry. Jahn's research into anomalous phenomena began over seven years ago, but it will be several more years before we know whether it has managed to progress much beyond previous attempts to bring scientific order into the field.

During the 130 year history of psychical research many other scholars and scientists initiated investigations of psychic phenomena with equally high hopes of taming the phenomena. One was the philosopher Henry Sidgwick who was the first president of the Society of Psychical Research founded in 1882. According to William James, Sidgwick and his colleagues "hoped that if the material were treated rigorously and, as far as possible, experimentally, objective truth would be elicited, and the subject rescued from sentimentalism on the one side and dogmatizing ignorance on the other. Like all founders, Sidgwick hoped for a certain promptitude of result; and I heard him say, the year before his death, that if anyone had told him at the outset that after twenty years he would be in the same identical state of doubt and balance that he started with, he would have deemed the prophecy incredible. It appeared impossible that the amount of handling evidence should bring so little finality of decision" [2].

James, who made this observation in his last article on psychical research in 1909, continued as follows [2]:

> My own experience has been similar to Sidgwick's. For twenty-five years I have been in touch with the literature of psychical research, and have had acquaintance with numerous "researchers." I have also spent a good many hours (though far fewer than I ought to have spent) in witnessing (or trying to witness) phenomena. Yet I am theoretically no "further" than I was at the beginning; and I confess that at times I have been tempted to believe that the Creator has eternally intended this department of nature to remain *baffling*, to prompt our curiosities and hopes and suspicions all in equal measure, so that, although ghosts and clairvoyances, and raps and messages from spirits, are always seeming to exist and can never be fully explained away, they also can never be susceptible of full corroboration.
>
> The peculiarity of the case is just that there are so many sources of possible deception in most of the observations that the whole lot of them *may* be worthless, and yet that in comparatively few cases can aught more fatal than this vague general possibility of error be pleaded against the record. Science, meanwhile needs something more than bare possibilities to build upon; so your genuinely scientific inquirer—I don't mean your ignoramous "scientist"—has to remain unsatisfied.

Some 67 years after James' final word on the matter, the philosopher Antony Flew summed up his 25 years of interest in parapsychology with remarkably similar sentiments [3]:

> My long-out-of-print first book was entitled, perhaps too rashly, *A New Approach to Psychical Research*. . . . When I reviewed the evidential situation at that time it seemed to me that there was too much evidence for one to dismiss. Honesty required some sort of continuing interest, even if a distant interest. On the other hand, it seemed to me then that there was no such thing as a reliably repeatable phenomenon in the area of parapsychology and that there was really almost nothing positive that could be pointed to with assurance. The really definite and decisive pieces of work seemed to be uniformly negative in their outcome.
>
> It is most depressing to have to say that the general situation a quarter of a century later still seems to me to be very much the same. An enormous amount of further work has been done. Perhaps more has been done in this latest period than in the whole previous history of the subject. Nevertheless, there is still no reliably repeatable phenomenon, no particular solid-rock positive cases. And yet there still is clearly too much there for us to dismiss the whole business.

Sidgwick was assessing the first 50 years of psychical research. James was evaluating the same period with another ten years or so added. Flew based his assessment on an additional 67 years of inquiry. Yet, all three agree that they could detect no progress. In each case after a quarter of a century of personal involvement, the investigator found the evidence for the paranormal just as inconclusive as it had been at the beginning. James openly concedes that *all* the claimed phenomena might be the result of self-deception or fraud. Yet he, and the other two philosophers, cannot quite shake the conviction that, despite all this inconclusiveness, "there might be something there."

Over this same span of history, the critics have consistently insisted that "there is nothing there." All the alleged phenomena of telepathy, clairvoyance, psychokinesis, levitation, spirit materialization, and premonitions can be accounted for in terms of fraud, self-delusion, and simple gullibility. The proponents have naturally resented such dismissals of their claims. They have argued that the critics have not fairly examined the evidence. They have accused the critics of attacking the weakest evidence and of ignoring the stronger and better supported evidence in favor of the paranormal.

Unfortunately, as any reading of the history of psychical research quickly reveals, the psychical researchers are correct in their appraisal of their critics. Too often, the major critics have attacked strawmen and have not dealt with the actual claims and evidence put forth by the more serious researchers. The fact that most of the criticism of the psychical research has been irrelevant and unfair, however, does not necessarily mean that the psychical researchers have a convincing case.

Indeed, the message that we get from Sidgwick, James, Flew, and Jahn is that the evidential base for psychic claims is very shaky at best. At most, these scholars, after carefully weighing all the evidence available to them, are claiming only that they cannot help feeling that, despite the inconsistencies and nonlawfulness of the data, that "there must be something there."

As will be discussed later in this paper, both the critics and the proponents subscribe to what I refer to as the False Dichotomy. When a scientist or scholar, after investigating possible psychic phenomena, concludes that the phenomena are real, the assumption is that either his conclusion is justified or he is delinquent in some serious way—being either incompetent or subject to some pathology. When the critic denies that the claim is justified, the proponent feels that his integrity or competence is being challenged. And the critic, sharing in this assumption, feels that he must show that the claimant is incompetent, gullible, or deficient in some serious way [4].

I consider this a False Dichotomy because competent and honest investigators can make serious judgmental errors when investigating new phenomena. Competence and expertise in any given field of endeavor is bounded. Cognitive

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

psychologists, historians of science, and sociologists of knowledge have been gathering data which demonstrate how thinking is guided by conceptual frameworks and paradigms within which the thinker operates. Successful scientific thinking, for example, is not successful because it operates according to abstract, formal rules of evidence. Rather, it succeeds because the thinker is guided by the often implicit rules and procedures inherent within the specific content and practices of the narrow field of specialization within which the problem is being pursued. These "heuristics" or guidelines for successful thinking are not foolproof and under changed circumstances they can trap the thinker into erroneous convictions. In other words, competence in a given scholarly or scientific discipline and high intelligence are no barriers to becoming trapped into asserting and defending erroneous positions.

In this paper, I agree with Sidgwick, James, Flew, and Jahn in the most general sense that "something" is indeed going on. However, I do not see any need to assume that this "something" has anything to do with the paranormal.

I think we should not lightly dismiss the fact that for 130 years some of our best scholars and scientists have seriously carried out psychical research and have become convinced that they have demonstrated the existence of a "psychic force" or a supernatural realm occupied by intelligent and superior beings. As far as I can tell, these proponents were competent scholars, sane, and highly intelligent. They made every apparent effort to employ what they believed to be objective and scientific standards in observing, recording, and reporting their findings.

Yet, as I will argue, contrary to Jahn's assessment, the total accumulation of 130 year's worth of psychical investigation has not produced any consistent evidence for paranormality that can withstand acceptable scientific scrutiny. What should be interesting for the scientific establishment is not that there is a case to be made for psychic phenomena, but rather that the majority of scientists who decided to seriously investigate *believed* that they had made such a case. How can it be that so many outstanding scientists, including several Nobel Prize winners, have convinced themselves that they have obtained solid, scientific evidence for paranormal phenomena?

If they are wrong, what has made them wrong? Does this suggest weaknesses or limitations of scientific method and training? And if these investigators have not actually encountered psychic phenomena, what is it that they have discovered?

I am not sure that I can provide satisfactory answers to these questions. But I believe that it will help to look at some selected cases in which investigators believed that they had obtained adequate scientific evidence for the reality of psychic phenomena. I will start at the beginning by describing the sort of evidence that convinced the first scientists who took psychical claims seriously. Even some contemporary parapsychologists believe these early scientists may have been wrong, but their cases are still worth examining because in them we will find many of the same issues and problems that characterize contemporary parapsychological research. These early psychic investigators tested spiritualistic mediums who were noted for their ability to produce powerful psychic phenomena such as levitations, materializations, and other physical feats.

Psychical research became transformed into what is now called parapsychology when the focus shifted, after the first half century of investigation, to the study of extrasensory perception and psychokinesis in ordinary individuals by means of standardized testing materials and procedures. I will examine what was, at the time, considered to be the most rigorous and successful application of this form of parapsychological research—the now notorious investigations by Soal on Shackleton and Mrs. Stewart. Again, the purpose is not to beat a dead horse but to abstract out principles and issues that still haunt contemporary parapsychology.

The card-guessing experiments begun by Rhine in the 1930s established the paradigm which dominated parapsychology for the next 40 years. New technology and interest in altered states resulted in departures from Rhine's paradigm beginning about 1970. Experiments with Random Event Generators, Remote Viewing, and the Ganzfeld technique have been the strongest contenders for providing parapsychology with its long-sought-for repeatable experiment. I will argue that a fair and objective assessment of this latest work strongly suggests that, like its predecessors, it still does not stand up to critical scrutiny.

Scientists and Psychics

The first major scientist to test experimentally a psychic claim was Michael Faraday in 1853. As will be described in more detail in the next section, Faraday concluded that the phenomena he had investigated, table-turning, had a normal explanation. Robert Hare, a major American chemist, at first agreed with Faraday's conclusion. But, then, after personal investigations of his own, changed his mind, and openly supported the claims of spiritualistic mediums. A decade later, Alfred Russel Wallace, the cofounder with Darwin of the theory of evolution by natural selection, and Sir William Crookes, the discoverer of thallium, astounded their scientific colleagues by openly endorsing paranormal claims. Wallace and Crookes, as had Hare, believed that their own inquiries had established scientific proof to support their paranormal claims.

Hare, Wallace, and Crookes were the first of a continuous succession of eminent scientists who have endorsed paranormal claims as a result of their experimental tests of alleged psychics. These scientists have established a tradition which has played a major role in the development of psychical research. The first half-century of psychical research consisted mainly of testing paranormal claims within this tradition. Beginning in the 1930s a second approach, experimental investigations according to standard protocols and using unselected subjects, became the dominant approach under the name of parapsychology. Today parapsychology includes both approaches.

In the first half of the present paper, I will focus on the first approach. The research of Sir William Crookes will be used as an example of this approach. In the second half of the paper, I will deal with the second approach. Again, I will use the research of a single investigator to bring out the more general issues and problems with the field of parapsychology. In both parts of the paper I will also briefly mention other investigators and lines of research which also bring out the same themes illustrated by the more detailed examples. Finally, I will briefly look at the contem-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

porary situation in parapsychology to argue that the concerns and difficulties that haunted the earlier investigations still persist.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table-Turning and Psychical Research

Modern spiritualism began when unaccountable raps were heard in the presence of two teen-age girls, Margaret and Kate Fox, in 1848. By using a code, the girls' mother was able to converse with the raps and concluded that they originated from the spirit of a peddler who had been murdered in the very house in which the Fox family then lived. Word of this miraculous communication spread quickly and soon a variety of means for communicating with the unseen spirits via "the spiritual telegraph" were developed in the United States and then spread to Europe. The individuals through whom the spirits produced their phenomena and communicated with mortals were called mediums. The mediums, at first, displayed phenomena such as rapping sounds, movements of tables and objects, playing of musical instruments by unseen agencies, and the occurrence of strange lights in the dark. Later, more elaborate phenomena were produced such as the levitation of objects or the medium; the disappearance or appearance of objects; the materialization of hands, faces, or even of complete spirit forms; spirit paintings and photographs; and written communications from the spirit world [5], [6].

By the early 1850s, table-turning (also called table-tilting or table-rapping) had become the rage both in the United States and in Europe. A group of individuals, usually called "sitters," would arrange themselves around a table with their hands resting flat upon the table-top. After an extended period of waiting a rap would be heard or the table would tilt up on one leg. Sometimes the table would sway and begin moving about the room, dragging the sitters along. On some occasions, sitters would claim that the table actually levitated off the floor under the conditions in which all hands were above the table. Reports even circulated that sometimes the table levitated when no hands were touching it. Table-turning was especially popular because it could occur with or without the presence of an acknowledged medium. Any group of individuals could get together and attempt to produce the phenomenon in the privacy of their own living room.

Table-turning plays an important role in the history of psychical research because it was what first attracted the attention of serious scientists to alleged paranormal phenomena [6]. The phenomenon had become so widespread in England by the summer of 1853 that several scientists decided to look into it. Although the prevailing explanation for the table's movements favored the agency of spirits, other explanations at the time were electricity, magnetism, "attraction," Reichenbach's Odyllic Force, and the rotation of the earth. Electricity, which in the public mind was then considered to be an occult and mystical force, was especially popular. Indeed, many spiritualists probably thought that the spirits operated by electricity.

In June 1853, a committee of four medical men held seances to investigate table-turning. They found that the table did not move at all when the sitters' attention was diverted and they had not formed common expectations about how the table should move. In another condition they found that the table would not move if half the sitters expected it to move to the right and the other half expected it to move to the left. "But when expectation was allowed free play, and especially if the direction of the probable movement was indicated beforehand, the table began to rotate after a few minutes, although no one of the sitters was conscious of exercising any effort at all. The conclusion formed was that the motion was due to muscular action, mostly exercised unconsciously" [6]. Other investigators came to similar conclusions.

But, by far, the most publicized and influential investigation was that by England's most renowned scientist: the physicist Michael Faraday. Faraday obtained subjects who were "very honorable" and who were also successful table-movers" [7]. Faraday found that he could obtain movements of the table in a given direction with just one subject sitting at his table in the laboratory. His first tests were designed to eliminated as explanations well-known forces such as magnetism and electricity. He demonstrated that substances such as sand-paper, millboard glue, glass, moist clay, tinfoil, cardboard, vulcanized rubber and wood did not interfere with the table-turning. He could find no traces of electrical or magnetic effects. "No form of experiment or mode of observation that I could devise gave me the slightest indication of any peculiar force. No attractions, or repulsion,... nor anything which could be referred to other than the mere mechanical pressure exerted inadvertently by the turner."

Although Faraday suspected that the sitter was unconsciously pushing the table in the desired direction, the sitter adamantly insisted that he was not the agency but, instead, was pulled in the expected direction by some force within the table. Faraday created some ingenious arrangements to see if the sitter's claim was true. He placed four or five pieces of slippery cardboard, one over the other, on the table top. The pieces were attached to one another by little pellets of a soft cement. The lowest piece was attached to a piece of sandpaper which rested on the table top. The edges of the sheets overlapped slightly, and on the under surface, Faraday drew a pencil line to indicate the position. The table-turner then placed his hands upon the upper card and waited for the table to move in the previously agreed upon direction (to the left). Faraday then examined the packet. It was easy to see by displacement of the parts of the line, that the hand had moved further than the table, and that the latter had lagged behind;—that the hand, in fact had pushed the upper card to the left and that the under cards and the table had followed and been dragged by it" [7].

In another arrangement, Faraday fixed an indicator to two boards on the table top such that if the sitter was pulled by the table the indicator would slope to the right but if the sitter pushed the table, the indicator would slope to the left. The table moved as before as long as the sitter could not see the indicator. But as soon as the sitter was able to watch the indicator, which gave him immediate feedback when his hands pushed in the expected direction, all movements of the table ceased. "But the most valuable effect of this test-apparatus ... is the corrective power it possesses over the mind of the table-turner. As soon as the index is placed before the most earnest, and they perceive—as in my presence they have always done—that it tells truly whether they are pressing downwards only or obliquely; then all effects of table-turning cease, even though the parties persevere, earnestly desiring motion, till they become weary and worn out. No prompting or checking of

 Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the hands is needed—[illegible] because the parties are made conscious of what they are really doing mechanically, and so are unable unwittingly to deceive themselves" [7].

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Faraday's investigation convinced several scientists that table-turning was the result of self-deception resulting from unconscious motor movements guided by expectation. His report is even credited with dampening the enthusiasm, for a few years, for spiritualism in England [6]. But several spiritualists and table-turners were not convinced by Faraday's arguments. And this brings up another issue that invariably accompanies the controversy over paranormal claims. Whenever a skeptic demonstrates how an alleged psychic phenomenon can be duplicated by mundane means, the claimant usually responds, "It's not the same thing!"

To many spiritualists and those who had witnessed table-turning, Faraday's explanation appeared hopelessly inadequate. Professional mediums, for example, while sitting at the table could provide meaningful answers by means of table-rapping to questions that sitters put to their assumed spirit communicators. In addition, the table often moved in a variety of ways which seemingly could not be explained by simple muscular pressure applied by the sitters. For example, the table often levitated above the floor with all the sitters' hands resting on the top surface. And some reports claimed that the table moved and levitated when no human was in contact with it.

Faraday's explanation dealt with only one important cause of the table-turning. He did not attempt to account for the various ways in which the table could be moved and levitated by trickery. Nor did he deal with the problem of the notorious unreliability of eyewitness testimony. Nor did he and his fellow skeptics realize that an abstract, even if correct, explanation of table-turning was impotent when matched against the personal and powerfully emotional experience of a sitter who has been converted during an actual table-turning session. These same limitations on any attempt to "explain away" an alleged paranormal event by a mundane account continue to provide loopholes whereby the proponent can maintain the reality of a paranormal claim.

Two striking illustrations of the power of the experience that "it is not the same thing," can be found in the conversions to spiritualism of the next two major scientists to investigate psychic phenomena. Both Robert Hare and Alfred Russel Wallace were familiar with Faraday's research and explanation when they first investigated spiritualistic phenomena by means of table-turning. And both were immediated convinced that their personal experiences could not be accounted for by Faraday's theory. In these instances, the forewarning, rather than serving to forearm, actually disarmed. And this, too, is a recurring theme in the history of psychical research.

Sir William Crookes

Faraday, the first major scientist to seriously investigate spiritualistic phenomena, concluded that self-deception was sufficient to explain what he observed. As a result, he remained skeptical and critical of all further claims of paranormal phenomena. Faraday's scientific colleagues were obviously grateful for his investigation and conclusions. But within the next two decades three other major scientists also investigated paranormal claims and concluded, con-[illegible] paranormal phenomena.

Robert Hare, the eminent American chemist, began his inquiry into spiritualistic phenomena in 1853 immediately after Faraday's investigation. Alfred Russel Wallace, the cofounder with Darwin of the theory of evolution by natural selection, initiated his investigations in 1865. And Sir William Crookes, the discoverer of thallium, began his investigations in 1869. All three had already achieved reputations as outstanding scientists before they surprised their scientific colleagues with their assertions of having witnessed psychic phenomena. Their colleagues were disturbed and puzzled by such assertions from obviously competent scientists. Their reactions, unfortunately, were not always rational and tended to make a confusing situation worse.

I believe it is important to try to understand how these otherwise competent scientists became convinced that they had acquired evidence sufficient to justify the belief in paranormal phenomena. The investigations of these scientists can be credited with the initiation of psychical research as a field with scientific aspirations. And many of the same issues of scientific justification of claims for the paranormal that we find in their work are still with us today.

Robert Hare was Professor Emeritus of Chemistry at the University of Pennsylvania and 72 years of age when circumstances conspired to launch him on a new career as a psychic investigator in 1853 [8]. Hare, the author of more than 150 scientific papers, had invented the oxy-hydrogen blowpipe which was the predecessor of today's welding torches [9]. According to Asimov, Hare was "one of the few strictly American products who in those days could be considered within hailing distance of the great European chemists" [10].

Both Hare and his critics took it for granted that a competent scientist could carry out observations and experiments on a variety of phenomena and, as a result, come to trustworthy and sound conclusions. Until he announced his conversion to the spiritualistic hypothesis, Hare's colleagues did not doubt his competence as an observer and experimenter. When he announced that he had not only experimentally verified paranormal phenomena, but had been communicating with the spirits of his departed relatives and also with George Washington, John Quincy Adams, Henry Clay, Benjamin Franklin, Byron, and Isaac Newton, this placed his incredulous colleagues in a quandary [8].

For half a century, the scientific world had accepted Hare's scientific papers and conclusions with respect and admiration. His scientific accomplishments were widely recognized and honored. But now this respected fellow scientist, by using apparently the same observational and experimental skills that had earned him his renown, was claiming to have demonstrated the reality of phenomena that scientists felt were just too preposterous to be true. Instead of examining Hare's arguments and evidence, his colleagues reacted emotionally and rejected his conclusions out of hand. Furthermore, they treated him as a traitor to the scientific enterprise and refused to allow him to present his case in the regular scientific forum.

From Hare's perspective this reaction was both unfair and unscientific. His arguments were being rejected without even being given a hearing. In his last few years he turned

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

away from his scientific colleagues and confined his social interactions entirely to his spiritualistic associates. From the perspective of the scientific establishment, Hare had suddenly gone insane or had suffered some other form of pathology. Here we see the False Dichotomy in action. And this same False Dichotomy will be found throughout the story of psychical research right up to the present.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Alfred Russel Wallace's conversion to spiritualism began in the same way that Hare's did—sitting at an animated table during a seance. Wallace's experience, just as Hare's did, convinced him that Faraday's explanation of the table's antics would not do. Unlike Hare, however, Wallace was not 72 and at the end of his career. Instead he was 42 years old and in the middle of a long and productive career. It had only been seven years earlier that Wallace had independently conceived the theory of evolution by natural selection, the very same theory that Darwin had been secretly working on for many years [11]–[13].

Critics have found it easy to dismiss the psychical evidence of Hare on the basis of old age and of Wallace on the assertion that, while he was a great naturalist and observer, he was not an experimenter [11]. Neither criticism can be applied, however, to William Crookes, who was the next great scientist to investigate and endorse the reality of paranormal phenomena. Crookes was generally acknowledged, even by many who opposed his psychic beliefs, as one of the preeminent chemists and physicists of his day. Crookes—the discoverer of thallium, inventor of the radiometer, developer of the Crookes tube, pioneer investigator of radiation effects, and a contributor to photography and other fields—was elected a Fellow of the Royal Society at age 31, was later knighted, and received just about every honor available to a scientist of his time.

When Crookes began attending seances with Mrs. Marshall (the same medium who helped convert Wallace) and J. J. Morse in 1869, he was 37 years of age. He had been very upset by the death of his youngest brother and apparently believed he had received spirit communications from him through the services of these mediums. In July 1870 Crookes announced his intention to conduct a scientific inquiry into spiritualistic phenomena. He wrote, "I prefer to enter upon the inquiry with no preconceived notions whatever as to what can or cannot be, but with all my senses alert and ready to convey information to the brain; believing, as I do, that we have by no means exhausted all human knowledge or fathomed the depths of all physical forces" [15].

Although most of the scientific community assumed that Crookes was undertaking the investigation as a skeptic, his biographer wrote, "But it is certain, at all events, that when in July 1870 Crookes, at the request, it is said of a London daily paper, announced his intention of 'investigating spiritualism, so-called,' he was already much inclined towards spiritualism. What he really intended to do was to furnish, if possible, a rigid scientific proof of the objectivity and genuineness of the 'physical phenomena of spiritualism,' so as to convert the scientific world at large and open a new era of human advancement" [16].

Crookes packed almost all his research into psychical phenomena into the four-year period 1870–1874 [17]. When he failed to sway his scientific colleagues—and as a result of bitter attacks by his critics, Crookes quietly dropped this work and devoted his scientific efforts from 1875 onwards to more mainstream subjects. But he never gave up his [illegible] ties with the field. In his final years, he began attending seances again and believed, near the end, that he had finally found proof of survival when he obtained a spirit photograph of his dead wife [15].

By today's standards, the investigations that come closest to being "scientific" were those that Crookes carried out with the celebrated medium Daniel Dunglas Home. Home is probably the most colorful and enigmatic psychic in the history of spiritualism [6], [9]. In one session, which took place at Crooke's home on May 31, 1871, Home held an accordian (which had just been purchased by Crookes for this occasion) by one end so that the end with the keys hung down towards the floor. The accordian was placed in a special cage under the table which just allowed Home's hand to be inserted to hold the accordian. Home's other hand was visible above the table. The individuals sitting on either side of Home could see his hand as well as the accordian in the wire cage. "Very soon the accordian was seen by those on each side to be moving about in a somewhat curious manner, but no sound was heard..." After putting the accordian down, Home picked it up again. This time several notes were heard. Crookes' assistant crawled under the table and said that he saw the accordian expanding and contracting, but Home's hand was quite still [15].

At the same session Crookes reported an experiment that he regarded as even "more striking, if possible, than the one with the accordian." A mahogany board, 3 ft long, with one end resting on a table and other end supported by a spring balance, was in a horizontal position. Home, while "sitting in a low easy-chair" placed the tips of his fingers lightly on the extreme end of the board which was resting on the table. "Almost immediately the pointer of the balance was seen to descend. After a few seconds it rose again. This movement was repeated several times, as if by successive waves of the Psychic Force. The end of the board was observed to oscillate slowly up and down during the experiment" [15].

To see if were possible to produce an effect on the spring balance by ordinary pressure, Crookes stood on the table and pressed one foot on the end of the board where Home had placed his fingers. By using the entire weight of his body (140 lb), Crookes was able to get the index to register at most 2 lb. Home had apparently achieved a maximum displacement of 6 lb.

Because of such results Crookes concluded that, "These experiments appear conclusively to establish the existence of a new force, in some unknown manner connected with the human organisation which for convenience may be called the Psychic Force" [15]. The skeptics were not convinced. They raised a variety of objections to the experiment measuring the movement of the board. Crookes thought some of the criticisms were unfair and irrelevant. But others he felt were reasonable and could be answered.

He repeated the experiment with additional controls. To avoid direct contact with the board, he altered the apparatus slightly in a manner that had previously been used by Robert Hare in some of his experiments. A bowl of water was placed on the end of the board not supported by the spring scale. Inside the bowl of water was lowered a "hemi-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

spherical copper vessel perforated with several holes at the bottom." The copper vessel was suspended from a large iron stand which w[illegible]tus. Home placed his fingers lightly in the water in the copper bowl. Presumably, this prevented him from having direct contact with the board. Yet, under these conditions Home managed to cause the other end of the board to sway up and down.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Finally, Home was removed a few feet away from the apparatus and his hands and legs were held. Even under these conditions, Crookes was able to record movements of the board, although the displacement was less the farther Home was from the apparatus. In further answer to critics, Crookes describes similar experiments carried out successfully by other researchers including Robert Hare. Crookes also got similar results using a lady who was not a professional medium in place of Home.

This series of experiments is by far the most impressive, from a scientific viewpoint, of any that Crookes conducted. Indeed, so far as I can tell, although these were among the very first serious attempts by a scientist to test a psychic, they have not been exceeded in degree of documentation and experimental sophistication during the subsequent 114 years. This is despite the fact that following Crookes' example, eminent scientists during almost every decade since Crookes' experiments have conducted tests of famous psychics.

The comments in the preceding paragraph should not be taken as an endorsement of Crookes' results. His experiments on the "Psychic Force" are superior *relative* to what has been reported by other scientists, including contemporary ones in their tests of psychic superstars. On an absolute scale of judgment the experiments still leave much to be desired. A major problem is documentation. Crookes omits many details which, from today's perspective at least, seem important in assessing what might have taken place.

Responding to the accusation that his witnesses were not reliable, Crookes wrote, "Accustomed as I am to have my word believed without witnesses, this is an argument which I cannot condescend to answer. All who know me and read my articles will, I hope, take it for granted that the *facts* I lay before them are correct, and that the experiments were honestly performed, with the single object of eliciting *the truth*" [15].

Here Crookes raises an important issue. When he reported finding a green line in a spectrum where one had never been reported, and followed this up with various analyses and controls to support the assertion that he must have discovered a new element (thallium), his scientific colleagues did not insist that he import skeptical witnesses, nor did they question his observations. The reported observation was made by using standard apparatus and recording procedures. The necessary controls and possibilities of error in such a context were well-known to workers in the field and it could be safely assumed that any trained chemist in this situation would behave according to both implicit and explicit rules.

But Crookes and his critics seriously err when they assume that similar confidence and trust can be placed in observations made in a field outside the investigator's training and one in which no standardization exists for instrumentation, making observations, instituting controls, recording the data, and reporting the results. The difficulties are compounded further when the observations are made, [illegible] materials, but of events involving humans who have a capacity to anticipate the experimenter's objectives and alter their behavior accordingly.

I recently discovered that Podmore, back in 1902, anticipated most of my reservations about Crookes' experiment on the movements of the balance [6]:

> The experiment as it stands, even without the modifications introduced later by Mr Crookes in deference to his scientific critics, seems, indeed, conclusive against the possibility of Home's affecting the balance by any pressure on his end of the board. But, tested by the canons laid down by Mr Crookes himself at the outset of his investigations, we shall find the conditions of the experiment defective in one important particular. Mr. Crookes had shown that it is the province of scientific investigation not merely to ascertain the reality of the alleged movements and measure their extent, but to establish their occurrence under conditions which render fraud impossible. In the passage quoted on page 183 it is implicitly recognised that such conditions are to be secured by eliminating the necessity for continuous observation on the part of the investigator. The proof of the thing done should depend upon something else than the mere observation of the experimenters, however skilled.
> Now in the experiment quoted these conditions were not fulfilled. On the contrary, we are expressly told that all present guarded Home's feet and hands. It is pertinent to point out that a duty for which the whole company were collectively responsible may well at times have been intermitted. Moreover, Dr. Huggins and Mr. Crookes had to watch the balance also, and Mr. Crookes had to take notes. Again, the experiment described was not the first of the kind; it occurred in the middle of a long series. It is indeed stated that Home was not familiar with the apparatus employed. But as similar apparatus had been employed probably at previous trials by Mr. Crookes himself, certainly by earlier investigators—amongst them Dr. Hare, with whose published writings on Spiritualism we cannot assume Home was unacquainted—the statement carries little weight. Further, a point of capital importance, there had apparently been many previous trials with various modifications of the apparatus and many failures; in Mr. Crookes' own words, "the experiments I have tried have been very numerous, but owing to our imperfect knowledge of the conditions which favour or oppose the manifestations of this force, to the apparently capricious manner in which it is exerted, and to the fact that Mr. Home himself is subject to unaccountable ebbs and flows of the force, it has but seldom happened that a result obtained on one occasion could be subsequently confirmed and tested with apparatus specially contrived for the purpose."
> The real significance of this statement is that Home—a practised conjurer, as we are entitled to assume—was in a position to dictate the conditions of the experiment. By the simple device of doing nothing when the conditions were unfavourable he could ensure that the light (gas in the present instance) was such and so placed, the apparatus so contrived, and the sitters so disposed, as to suit his purpose, and that in the actual experiment the attention of the investigators would necessarily be concentrated on the wrong points. Under such conditions, as ordinary experience shows, and as the experiments described in the last chapter have abundantly demonstrated, five untrained observers are no match for one clever conjurer.

Podmore is referring, in the last sentence, to the dramatic experiments on eye-witness testimony conducted by S. J. Davey [18]. Davey had been converted to a belief in spiritualistic phenomena by the slate-writing demonstrations of the medium Henry Slade. Subsequently, Davey

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

accidently discovered that Slade had employed trickery to produce some of the phenomena. Davey practiced until he felt he could accomplish all of Slade's feats by trickery and misdirection. He then conducted his well-rehearsed seance for several groups of sitters, including many who had witnessed and testified to the reality of spiritualistic phenomena. Immediately after each seance, Davey had the sitters write out in detail all that they could remember having happened during his seance. The findings were striking and very disturbing to believers. None of the sitters had suspected Davey of using trickery. Sitters consistently omitted crucial details, added others, changed the order of events, and otherwise supplied reports which would make it impossible for any reader to account for what was described by normal means.

Podmore has much more to say about this experiment. His reference to "untrained" observers is not meant to question Crookes' scientific competence. "But his previous training did not necessarily render him better qualified to deal with problems differing widely from those presented in the laboratory. To put it bluntly, if Home was a conjurer, Mr. Crookes was probably in no better position for detecting the sleight-of-hand than any other man his equal in intelligence and native acuteness of sense. Possibly even in a worse position; for it may be argued that his previous training would prepare the way for Home's efforts to concentrate attention on the mechanical apparatus, and thus divert it from the seemingly irrelevant movements by which it may be conjectured the conjurer's end was attained."

Finally, Podmore points out ways in which the report is incomplete. He then speculates about one possible way Home might have tricked Crookes. He describes a scenario in which Home could have employed a thread which he attached to the apparatus, probably the hook of the scale. Some further points could be mentioned such as the fact that Crooke's unpublished notes suggest that the experiment was much more informal and involved many more distractions than the published version indicates [15].

Crookes held many seances not only with Home but with almost every major spiritualistic medium who was in England during the years 1869 through 1875. He reported having observed a variety of phenomena which he argued could not have been produced by normal means: movement of heavy bodies with contact but without mechanical exertion; raps and other sounds; the alteration of weights of bodies; movements of heavy substances at a distance from the medium; the rising of tables and chairs off the ground, without contact of any person; the levitation of human beings; the appearance of hands, either self-luminous or visible by ordinary light; direct writing; and phantom forms and faces [18]. His documentation for such phenomena, however, falls far short of what he has supplied us for the movements of the balance.

As was the case with Hare and Wallace, Crookes was bitterly attacked for his views. The eminent physiologist, William Carpenter, lead the opposition. Carpenter openly questioned Crookes' competence as a scientist, wrongly stated that Crookes' election to the Royal Society had been questionable, and made several other unwarranted insults [16], [17]. Like Wallace, Crookes tried to get his scientific colleagues and critics to witness his experiments with Home and other psychics. But none of them accepted his invitations.

Hare, Wallace, and Crookes were the first of many eminent scientists who have investigated and endorsed psychics. Their work inspired many later scientists to also take time away from their regular scientific activities to investigate the paranormal claims of mediums or self-professed psychics [4], [19]–[29]. Yet, I suspect that many parapsychologists will object to using the work of these psychic investigators as part of a general evaluation and critique of parapsychology. The objection would be based on two arguments.

Today, most parapsychologists would not include the reports of Hare, Wallace, and Crookes in their case for the reality of psi (the current term to refer to extrasensory perception and psychokinesis). And, secondly, even the reports by more recent scientists on psychics do not form part of the primary database of parapsychology. Instead, today's parapsychologists want to base their argument on evidence emerging from laboratory experiments with unselected subjects and which use standardized tasks.

However, I believe there are good reasons for focussing on these early investigators:

1) At the time they were reported, these investigations were considered to be the strongest evidence for the paranormal. From 1850 to 1866 Hare's research constituted practically the entire "scientific" case upon which proponents could base their claims. From 1870 until the founding of the Society of Psychical Research in 1882 it was the work of Crookes and Wallace that proponents put forth as the best scientific justification for their paranormal claims.

2) The psychical research of these three eminent scientists served as the model for all later investigations of psychics by scientists. Although sometimes the latest technological developments are brought into the investigations, no change in approach or improvements in methodology for such investigations has occurred during the 130 years since Hare first reported his findings [23]. In terms of adequacy of documentation, for example, it is difficult to find any improvement over Crookes' reports on his experiments with Home in the subsequent accounts by such psychic investigators as Richet, Barrett, Lodge, Lombroso, Zoellner, Eisenbud, Targ, Puthoff, Hasted, and the many others.

3) The work of this early trio served as an important impetus for the subsequent founding of the Society for Psychical Research in 1882. In his presidential address to the first general meeting of The Society for Psychical Research on July 17, 1882, Henry Sidgwick went out of his way to acknowledge the importance and evidential value of the work of these pioneer researchers [30]:

> I say that important evidence has been accumulated; and here I should like to answer a criticism that I have privately heard which tends to place the work of our Society in a rather invidious aspect. It is supposed that we throw aside *en bloc* the results of previous inquiries as untrustworthy, and arrogate to ourselves a superior knowledge of scientific method or intrinsically greater trustworthiness—that we hope to be believed, whatever conclusions we may come to, by the scientific world, though previous inquirers have been uniformly distrusted. Certainly I am conscious of making no assumption of this kind. I do not presume to suppose that I could produce evidence better in quality than much that has been laid before the world by writers of indubitable scientific repute—men like Mr. Crookes, Mr. Wallace, and the late Professor de Morgan. But it is clear that from what I

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

have defined as the aim of the Society, however good some of its evidence may be in quality, we require a great deal more of it. I do not dispute — it is not now time to dispute —with any individual who holds that reasonable persons,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

who have looked carefully into the evidence that has been so far obtained, ought to be convinced by that evidence; but the educated world, including many who have given much time and thought to this subject, are not yet convinced, and therefore we want more evidence.

Sidgwick makes it clear that he and the other founders of the Society for Psychical Research consider the findings of Wallace and Crookes as scientifically sound. Sidgwick has no doubt that Wallace's and Crookes' reports *should* convince reasonable members of the scientific community. But he pragmatically makes the distinction between what *should* and what *will* convince the critics. "What I mean by *sufficient evidence* is evidence that will convince the scientific world, and for that we obviously require a good deal more than we have so far obtained" [30]. In other words, Sidgwick does not aspire to improve the quality of the preceding scientific investigators. Rather he wants to acquire more of the same quality.

4) The investigations of these original psychical researchers bring out many of the same issues of evidence, testimony, and proof that still characterize current controversies in parapsychology. Unfortunately, not much in the way of further clarification or resolution of these issues has occurred since their efforts first stimulated the debate. I have already mentioned some of these issues in my discussions of the individual cases.

Many of the issues involve the problem of competency. To what extent, for example, does competency in one branch of inquiry transfer, if at all, to a different branch? Can a scientist, no matter how competent and well-intentioned, initiate an inquiry into a previously unstructured and unstandardized area and single-handedly produce results which bear the same scientific status as the results he has produced in his original area of expertise? Elsewhere, I have given by reasons for answering this question in the negative [23].

One important issue is perhaps worth bringing up at this point. The scientists who have defended the trustworthiness of their psychical research have typically insisted that the observations and evidence of their reports of psychic happenings do not differ in quality from that which characterizes their more orthodox investigations.

Yet, at the same time, these same investigators acknowledge an important difference between their inquiries into physics and biology and their investigations of psychics. Hare, Wallace, and Crookes, as well as the later psychical researchers insisted that the psychics being tested must be treated with proper respect and concern for their feelings. If the investigator is overly skeptical or otherwise betrays distrust of the alleged psychic this could adversely affect the paranormal performance. Thus these scientists try to convey the impression that they conduct their tests using every precaution against fraud and deception, but at the same time making sure not to take any step or include any condition that meets with the disapproval of the alleged psychic. Skeptics such as myself, who have both experience in conducting experiments with humans and have been trained in conjuring, believe this is an impossible task. The twin goals of preventing trickery on the part of the alleged psychic and of ensuring that this same person will be satisfied with all the experimental arrangements are mutually incompatible.

But scientists who have testified to the paranormal powers of their subjects confidently insist they have simultaneously achieved both goals. A contemporary version of this theme has been eloquently put forth by a group of scientists, including two of England's outstanding physicists, in describing their experiments on the psychokinetic powers of Uri Geller [31]:

> We have come to realize that in certain ways the traditional ideal of the completely impersonal approach of the natural sciences to experimentation will not be adequate in this domain. Rather, there is a personal aspect that has to be taken into account in a way that is somewhat similar to that needed in the disciplines of psychology and medicine. This does not mean, of course, that is not possible to establish facts on which we can count securely. Rather, it means that we have to be sensitive and observant, to discover what is a right approach, which will properly allow for the subjective element and yet permit us to draw reliable inferences. One of the first things that reveals itself as one observes is that psychokinetic phenomena cannot in general be produced unless *all* who participate are in a relaxed state. A feeling of tension, fear, or hostility on the part of any of those present generally communicates itself to the whole group. The entire process goes most easily when all those present actively want things to work well. In addition, matters seem to be greatly facilitated when the experimental arrangement is aesthetically or imaginatively appealing to the person with apparent psychokinetic powers.
>
> We have found also that it is generally difficult to produce a predetermined set of phenomena. Although this may sometimes be done, what happens is often surprising and unexpected. We have observed that the attempt to concentrate strongly in order to obtain a desired result (e.g., the bending of a piece of metal) tends to interfere with the relaxed state of mind needed to produce such phenomena . . . . Indeed, we have sometimes found it useful at this stage to talk of, or think about, something not closely related to what is happening, so as to decrease the tendency to excessive conscious concentration on the intended aim of the experiment . . . .
>
> In the study of psychokinetic phenomena, such conditions are much more important than in the natural sciences, because the person who produces these phenomena is not an instrument or a machine. Any attempt to treat him as such will almost certainly lead to failure. Rather, he must be considered to be one of the group, actively cooperating in the experiment, and not a "subject" whose behavior is to be observed "from the outside" in as cold and impersonal manner as possible . . . .
>
> In such research an attitude of mutual trust and confidence is needed; we should not treat the person with psychokinetic powers as an "object" to be observed with suspicion. Instead, as indicated earlier, we have to look on him as one who is working with us. Consider how difficult it would be to do a physical experiment if each person were constantly watching his colleagues to be sure that they did not trick him. How, then, are we to avoid the possibility of being tricked? It should be possible to design experimental arrangements that are beyond any reasonable possibility of trickery, and that magicians will generally acknowledge to be so. In the first stages of our work we did, in fact, present Mr. Geller with several such arrangements, but these proved to be aesthetically unappealing to him. From our early failures, we learned that Mr. Geller worked best when presented with many possible objects, all together on a metal surface; at least one of these objects might appeal to him sufficiently to stimulate his energies . . . .
>
> Nevertheless, we realize that conditions such as we have described in this paper are just those in which a conjuring trick may easily be carried out. We understand also that we are not conjuring experts, so if there should be an intention to deceive, we may be as readily fooled as any person.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

> Moreover, there has been a great deal of public criticism, in which the possibility of such tricks has been strongly suggested. For this reason it has often been proposed that a skilled magician should be present to help to see that there will be no possibility of deception. It is in the nature of the case, however, that no such assurance can actually be given. For a skilled magician is able to exploit each new situation as it arises in a different and generally unpredictable way.... In principle, we would welcome help of this kind in decreasing the possibility of deception. It has been our observation, however, that magicians are often hostile to the whole purpose of this sort of investigation, so they tend to bring about an atmosphere of tension in which little or nothing can be done. Indeed, even if some magicians who were found who were not disposed in this way, it does not follow that their testimony will convince those who are hostile, since the latter can always suppose that new tricks were involved, beyond the capacity of those particular magicians to see through them. Because of all of this, it seems unlikely that significant progress towards clearing up this particular question could be made by actually having magicians present at the sessions, though we have found it useful to have their help in a consultative capacity.... We recognize that there is a genuine difficulty in obtaining an adequate answer to criticisms concerning the possibility of tricks, and that a certain healthy skepticism or doubt on the part of the reader may be appropriate at this point.... However, we believe that our approach can adequately meet this situation.

These investigators close this discussion of the difficulties of carrying out such research with an optimistic prognosis, "We feel that if similar sessions continue to be held, instances of this kind might accumulate, and there will be no room for reasonable doubt that some new process is involved here, which cannot be accounted for, or explained, in terms of the laws of physics at present known. Indeed, we already feel that we have very nearly reached this point." These hopeful words were written in 1975. Neither they nor other scientists have yet managed to present scientific evidence that Uri Geller or his many imitators can bend metal paranormally. Although at least one major physicist continues his investigations of paranormal metal bending [20], a decade of research on Uri Geller by scientists who adhered to the advice of treating the metal-bender as a respected colleague and catering to his aesthetic sensibilities has only succeeded to demonstrate that Geller can bend metal under conditions which allow him to do it by cheating [21].

Hare, Wallace, and Crookes, as well as subsequent psychic researchers, insisted they had guarded against the possibility of trickery while, at the same time, acknowledging the necessity to treat their subjects in the special way described by Hasted *et al.* Unfortunately, as Hasted *et al.* concede, this special treatment increases the difficulties of preventing deception. But, like their predecessors in psychical research, they express confidence that their scientific skills can overcome the difficulty. In fact, the suggested procedure gives the alleged psychic veto power over any arrangement that impedes trickery and also supplies a ready excuse for not producing phenomena when the dangers of detection suddenly seem too high. The conditions which the scientists report as ideal for the production of psychical phenomena are just those that are also ideal for the production of the same phenomena by trickery.

5) As already discussed, Hare, Wallace, and Crookes were bitterly attacked by their skeptical scientific colleagues. And the same sorts of attacks and defenses have characterized subsequent cases. Both critics and defenders still implicitly subscribe to the same False Dichotomy. And both the critics and the defenders, in different ways, do not emerge as rational, objective, scientific or otherwise admirable in their exchanges. Worse, no lessons from the past seem to have either been learned or carried over to the current controversies. If the critical exchanges had been more constructive and rational at the time of Hare, Wallace, and Crookes, today we might be closer to understanding what was really going on to make such eminent scientists put forth such seemingly outrageous claims.

Hare, Wallace, and Crookes had no success in inducing their critics to come and examine the evidence for themselves. It is possible that if Huxley and Carpenter had accepted Wallace's invitation to attend at least six seances, no phenomena would have taken place. On the other hand, it would be useful to have the accounts of such skeptical observers before us if, say, Miss Nichol did produce the flowers in their presence. And it certainly would have helped if Carpenter and Stokes had accepted Crookes' invitation to watch his experiments with Home and the balance.

The Creery Sisters

For its first 30 years, psychical research consisted of individual and uncoordinated investigations by scholars or scientists such as Hare, Wallace, and Crookes. During this period some feeble and unsuccessful attempts were made to form research societies and coordinate the research [32]. The first successful attempt to institutionalize psychical research was the founding of the Society for Psychical Research in London in 1882. Four of the principal leaders of this society—the philosopher Henry Sidgwick, the physicist William Barrett, the literary scholar Edmund Gurney, and the classicist Frederic Myers—had been encouraged, in addition to their own investigations of telepathy and mediums, by the research of such scientists as Wallace and Crookes. The founders of the Society clearly believed that they possessed solid scientific evidence for the reality of thought-transference. At the first general meeting of the Society in London on July 17, 1882, Henry Sidgwick ended his presidential address with the following words [30]:

> We must drive the objector into the position of being forced either to admit the phenomena as inexplicable, at least by him, or to accuse the investigators either of lying or cheating or a blindness or forgetfulness incompatible with any intellectual condition except absolute idiocy. I am glad to say that this result, in my opinion, has been satisfactorily attained in the investigation of thought-reading. Professor Barrett will now bring before you a report which I hope will be only the first of a long series of similar reports which may have reached the same point of conclusiveness.

Before looking at the experimental results whose "conclusiveness" Sidgwick believes is beyond reasonable doubt, I would like to call the reader's attention to the use of the False Dichotomy in Sidgwick's strategy. The goal is to report evidence that is so compelling that the critic either has to admit that psychic phenomena have been demonstrated or that the investigator is deliberately lying, afflicted with a pathological condition, or incredibly incompetent. Sidgwick does not allow for the possibility that an investigator could be competent, honest, sane, and intelligent, and still wrongly report what he believes to be "conclusive"

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

evidence for the paranormal. Unfortunately, as seen in the cases of Hare, Wallace, and Crookes and as typifies succeeding the cases, the critics, in responding to paranormal claims, have implicitly accepted the False Dichotomy. When confronted with paranormal claims by otherwise competent investigators, many critics have taken the bait and have tried to discredit the offending investigator by questioning his competence, insinuating fraud, or suggesting pathology.

The "conclusive" evidence with which Sidgwick wanted to confront the objector came from a series of experiments on thought-transference conducted by his colleagues William Barrett, Edmund Gurney, and Frederic Myers [33]. The investigators introduced this series as follows [33]:

> In the correspondence we have received there were two cases which seemed, upon inquiry, to be free from any *prima facie* objections, and apparently indicative of true thought-reading. One of these cases is given in the Appendix... but as we cannot from personal observation testify to the conditions under which the trials were made, we simply leave it aside. The other case was that of a family in Derbyshire, with whom we have had the opportunity of frequent and prolonged trials.
> Our informant was Mr. Creery, a clergyman of unblemished character, and whose integrity indeed has, it so happens, been exceptionally tested. He has a family of five girls, ranging now between the ages of ten and seventeen, all thoroughly healthy, as free as possible from morbid or hysterical symptoms, and in manner perfectly simple and childlike. The father stated that any one of these children (except the youngest), as well as a young servant-girl who had lived with the family for two years, was frequently able to designate correctly, without contact or sign, a card or other object fixed on in the child's absence. During the year which has elapsed since we first heard of this family, seven visits, mostly of several days' duration, have been paid to the town where they live, by ourselves and several scientific friends, and on these occasions daily experiments have been made.

The preceding quotation was taken from the "First Report on Thought-Reading" which was read at the first meeting of the Society. Several more experiments were conducted with the Creery sisters and the results included in the second and third reports [34], [35]. Notice the emphasis placed upon Reverend Creery's "unblemished character" and integrity. Within the Victorian society of Sidgwick and his colleagues this emphasis on character had a special significance. According to Nicol, many flaws in the investigative reports of the Society were due to "a *double standard* of evidence."

> The Society's double standard of evidence arose in the following way. The Society's leaders were members of the middle and upper middle strata of society. When faced with the problem of estimating the value of evidence, they divided the world into two classes: (a) Members of their own class (Ladies and Gentlemen in the Victorian sense) whom they tended to treat trustingly; (b) Members of the lower classes, whom for brevity we may call the Peasants: them they treated with suspicion [36].

The experiments with the Creery sisters were all variants of the popular Victorian pastime known as the "willing game" [37].

> The game admits of many variations, but is usually played somewhat as follows. One of the party, generally a lady, leaves the room, and the rest determine on something which she is able to do on her return—as to take a flower from some specified vase, or to strike some specified note on the piano. She is then recalled, and one or more of the 'willers' place their hands lightly on her shoulders. Sometimes nothing happens; sometimes she strays vaguely about; sometimes she moves to the right part of the room and does the thing, or something like the thing, which she has been willed to do. Nothing could at first sight look less like a promising starting-point for a new branch of scientific inquiry.

Barrett, Gurney, and Myers go to great lengths to assure their readers that they are aware of the many nonparanormal ways in which information from the senders can be communicated to the percipient. Subtle unconscious pushes by the "willer," for example, can guide the percipient to the correct place. And there is always the possibility of secret codes being employed [33], [37]. Nevertheless, they relate incidents from their own experience with the game which they believe cannot be handled by such obvious explanations.

In their typical experimental procedure, one child would be selected to leave the room. When she was out of the room, the remaining participants would select a playing card or write down a number or name. "On re-entering she stood—sometimes turned by us with her face to the wall, oftener with her eyes directed towards the ground, and usually close to us and remote from her family—for a period of silence varying from a few seconds to a minute, till she called out to us some number, card, or whatever it might be" [33]. Before leaving the room, the child was always informed of the general category, such as playing cards, from which the target item was to be chosen.

The authors obviously felt that their knowledge of the various ways that inadvertent and deliberate signaling of the percipient could occur somehow made them immune from such errors. As an added precaution, however, they conducted several trials either in which members of the family were absent or in which only the experimenters knew the chosen object (unfortunately they do not distinguish among trials on which only the experimenters were informed of the target but the family was present and trials on which only the experimenters were present). The investigators claim that keeping the family uninformed did not appreciably lower the proportion of above-chance correct guesses.

The results were quite striking. Looking only at the results of those trials on which members of the Committee alone knew the card or number selected, the investigators summarize their findings as follows [35]:

> 260 Experiments made with playing cards; the first responses gave 1 quite right in 9 trials; whereas the responses, if pure chance, would be 1 quite right in 52 trials. 79 Experiments made with numbers of two figures; the first responses gave 1 quite right in 9 trials; whereas the responses, if pure chance, would be 1 quite right in 90 trials.

The experimenters also summarize the results of the much larger number of trials in which the family members were not excluded. Two points are worth noting about the results reported above. By ordinary statistical criteria the odds against such an outcome being due just to chance are enormous. But the calculation of such odds assumes, that in the absence of telepathy, we know the expected value and distribution of hits. The way experimenters can ensure the appropriate conditions for the application of the statistical tests is to include careful procedures for randomizing the

 Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

targets on each trial such that each target has an equal chance of being selected and that the selected object on a given trial is independent of the selection on the next. But

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

nowhere in the three reports do we find any mention of how the playing card or number was chosen on each trial. We do not know if the deck was shuffled even once, let alone between trials. The number selection is even more disturbing because if, as seems to be the case, a committee member simply thought of any two digit number that came to mind, we know that some numbers are much more likely than others. And the same few numbers that are favored by the sender are likely to be those that come to the mind of the percipient. These most probable numbers, known as "mental habits" in the older literature, are called "population stereotypes" by Marks and Kammann [25].

The second peculiarity, which was noted by Coover, is that the proportion of successful hits in these experiments seems to be independent of the chance probability [38]. Thus the hit rate is 1 out 9 trials regardless of whether cards or numbers are being guessed. To Coover this suggest the use of a code rather than the imperfect transmission of psychic signals.

As already indicated, the founders of the Society for Psychical Research believed that, with the experimental results on the Creery sisters, they had finally succeeded in scientifically establishing telepathy as a valid phenomenon. As just one example of the importance attached to these experiments, Gurney's statement in the Society's first major monograph, *Phantasms of the Living* [39] can be cited:

> I have dwelt at some length on our series of trials with the members of the Creery family, as it is to those trials that we owe our own conviction of the possibility of genuine thought-transference between persons in a normal state.

Despite this confidence in the conclusiveness of the Creery experiments, critics quickly pointed out perceived flaws [38], [40], [41]. It was charged that the authors grossly underestimated the extent to which sophisticated coding could transpire between the girls in the experimental situation. The critics also suggested that the experimenters were naive in assuming that they could prevent inadvertent cueing just by being aware of the possibility.

Concerning the trials in which only the investigators knew the chosen object, the critics complained about inadequate documentation. The experimenters never state how the card or object was chosen; whether the members of the family were present during the selection (even though they were presumably kept ignorant of the choice); whose deck of playing cards was used; and so forth.

As can be seen, even on this brief account, we encounter a number of the issues that characterized earlier psychical research. The investigators assume that to be forewarned is to be forearmed. For example, they devote six pages of their first report to a discussion of the various types of errors, which if not excluded, could invalidate their research [33]. The purpose is to assure the reader that because they are keenly aware of the possibilities of such errors they could not have occurred. As previously mentioned, one way the investigators tried to preclude giving the girl any involuntary muscular cue was simply for the investigator to be consciously aware of such a possibility and consciously prevent himself from displaying such cues. Not only is such a precaution useless [42], but it was unnecessary since one could more directly prevent unwitting bodily cues by simply screening those who know the target from the percipient. This tendency to substitute *plausible* (to the investigator) reasons for discounting a possible source of error for actual experimental controls to guard against the error characterizes psychical research from its inception to the present.

A second theme is that prior experience in investigating paranormal claims automatically qualifies one as an expert who can be trusted not to make mistakes or be susceptible to trickery in future situations. This theme is closely related to the False Dichotomy issue.

The report on the Creery sisters also illustrates another recurring theme in psychical research—the *Patchwork Quilt Fallacy*. As Giere points out, the "patchwork quilt fallacy" gets its name because, "The hypothesis, initial conditions, and auxiliary assumptions are pieced together in such a way that they logically imply the known facts" [43]. Telepathy or psi always seems to be just that mysterious phenomenon that produced all the peculiar patterns that we happened to observe in our data. On some days the Creery sisters performed no better than chance. This variability among days became, in the minds of the investigators, a property of the phenomenon [35]:

> It may be noted that the power of these children, collectively or separately, gradually diminished during these months, so that at the end of 1882 they could not do, under the easiest conditions, what they could do under the most stringent in 1881. This gradual decline of power seemed quite independent of the tests applied, and resembled the disappearance of a transitory pathological condition, being the very opposite of what might have been expected from a growing proficiency in code-communication.

The fact that alleged psychics inevitably seem to lose their powers under continued investigation has become known as the "decline effect," which can occur in a variety of patterns and guises. Gurney and his colleagues propose the decline as additional support for the genuineness of the telepathy because it is not what might be expected if the girls were becoming more proficient in using a code. The cynic, of course, views this decline in the just the opposite way. Presumably the investigators are also becoming more proficient in knowing what to look for, especially in the face of continuing criticism, and, as a result, they have made it more difficult for the girls to get away with their tricks.

As it turns out the investigators later caught the girls cheating. The girls, at least on this occasion, had used a simple code. This brings up an additional theme in psychical research which we might, for short, label the *Problem of the Dirty Test Tube*. Gurney revealed the deception in a brief note which appeared in the *Proceedings of the Society for Psychical Research* in 1888 [44]. Hall thinks it is very significant that Gurney's fellow investigators did not sign this revelation [41].

In the note, Gurney reminds his readers "that the earliest experiments in Thought-transference described in the Society's *Proceedings* were made with some sisters of the name of Creery. The important experiments were, of course, those in which the 'agency' was confined to one or more of the investigating Committee.... But though stress was never laid on any trials where a chance of collusion was afforded by one or more of the sisters sharing in the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

'agency,' nevertheless some results contai… …nder such condit… …ssi… therefore, to state that in a series of experim… …with cards, recently made at Cambridge, two of the s… …acting as 'agent' and 'percipient,' were detected in th… …a code of signals; and a third has confessed to a c… …mount of signalling in the earlier series to which ref… …has been made" [44]. Gurney then describes both … visual and auditory codes used by the girls. He conti… …s follows [44]:

> The use of the visual code was very gratuitous … the part of the sisters, since it had been explained to the… …that we did not attach any scientific value to the experim… …s in which they acted as agent and percipient in sight of … other, the possibility of success under these condition… …aving been abundantly proved. The object of our experi… …nts at Cambridge on this occasion was, if possible, to … ngthen the evidence for Thought-transference (1) when … members of the family were aware of the thing to be gu… …ed, and (2) when the sister acting as agent was in a differ… room from the one acting as percipient. The experiment… …which the codes were used were intended merely as a… …sement and encouragement with a view to increase the cha… …e of success in the more difficult ones—which were all co… …lete failures. The account which was given as to the earlie… …experiments, conducted under similar conditions, is that sig… …s were very rarely used; and not on specially successful occasions, but on occasions of failure, when it was feared that visitors would be disappointed. But of course the rec… …nt detection must throw discredit on the results of all pre… …ous trials in which one or more of the sisters shared in the agency. How far the proved willingness to deceive can be held to affect the experiments on which we relied, where collision was excluded, must of course depend on the degree of stringency of the precautions taken against trickery of other sorts —as to which every reader will form his own opinion.

This manner of treating the discovery of cheating illustrates a number of interwoven themes. The finding of a "dirty test tube" ordinarily implies that all the results of the experiment are brought into question. Gurney argues that only those results clearly attached to the "dirty test tube" should be discarded. Since the girls could not have used their code, in his judgment, in those trials in which only investigators knew the chosen object, those trials still retain their evidential value. Related to this is what the early psychical researchers called the problem of "mixed mediumship." Psychics and mediums are under constant pressure to produce results, yet they have little direct control over their fickle powers. Therefore, in order not to disappoint their followers or from fear of losing the attention that goes with mediumship, they learn to supplement their real powers with tricks to simulate the phenomena. Still another variant of this exploits the apparent fact that many mediums and psychics are apparently in a trance or altered state when performing. In such a state they are highly suggestible and behave in ways expected of them. If skeptics are among the onlookers, they will sometimes cheat because this is what is expected of them. The onus for the consequent cheating is by this means placed upon the skeptic rather than the cheater.

The dirty test tube problem has been with psychical research from its beginning and, as we will see, is still very much a part of the contemporary scene. The medium Eusapia Palladino's long career was noteworthy for the number of times she was caught cheating. She readily acknowledged that she would cheat if the investigators gave her the opportunity. Despite this record of cheati… …including some of toda… leaders in the field, have no doubt that on many oth… occasions she displayed true paranormal powers [19]. … the contemporary scene, parapsychologists are willing … admit that the controversial metal-bender Uri Geller oft… cheats, but that, on occasion, he exhibits real paranorm… powers [45]. And parapsychologists blamed me, rather th… Geller, for the fact that Geller cheated in my presen… because, as they put it, I did not impose sufficiently str… gent conditions to prevent him from cheating [22].

Despite this attempt to save some of the evidence fr… the Creery experiments, the leaders of the Society for P… chical Research quietly removed the experiment from th… evidential database. But Sir William Barrett refused to … along with this demoting of the experiment. According … Gauld, this incident sparked dissension between Barr… and the other founders [32].

> Barrett had been the first to experiment with these girls, and they were his special proteges... Barrett would never agree that the later and crude cheating invalidated all the earlier results; he considered that his 187… experiments, together with his experiments with the Creerys had established his claim to be the discoverer of thought-transference, and he remained bitter towards the Sidgwicks for the rest of his life.

Not only did Barrett continue to defend the evide… value of the Creery experiments, but so did later … psychologists. In his classic monograph of 1934 on … *Sensory Perception*, J. B. Rhine included th… experim… among the most evidential of the early research. "… whole the early experiments in E.S.P. were admirab… ducted...as one would expect from the array of … impressive names connected with them. The exper… with the Creery sisters, for instance, were conduc… Professors William Barrett, Henry Sidgwick and … Stewart, by Mrs. Henry Sidgwick, Frederic Myers, Ed… Gurney and Frank Podmore.... In all this work the… were sufficiently striking to leave no doubt as to the … sion of the hypothesis of chance" [46].

Despite these attempts to salvage something from … Creery experiments, I believe it is fair to say that today … experiments are not part of the case that parapsycholo… would make in support of psi. Indeed, my perusal of sev… contemporary books and histories of parapsychology … dicates that the experiments are rarely, if ever, mention…

The same fate befell the very next major experiment … telepathy conducted by the same investigators. In th… "Second Report on Thought-Transference," Gurney and … colleagues describe the first of their experimental findi… in which two young men, Smith and Blackburn, were … parently able to communicate telepathically under con… tions that prevented normal communication. If anythi… the investigators placed even more reliance upon th… later experiments than in those with the Creery sisters.

As was the case with the Creery sisters, Smith and Bla… burn soon lost their powers. Smith was then hired by … Society to assist in the conduct of several successful te… pathic experiments. In 1908, Blackburn, thinking that Sm… was dead, publicly confessed as to how he and Smith … tricked the investigators during the experiments. Sm… who was very much alive and still employed by the Soci… denied the charges. In the ensuing debate, the Socie…

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

leaders defended Smith. Good accounts of this amazing incident can be found in [38] and [41]. Today, the Smith-Blackburn experiments are no longer considered part of the parapsychological case for psi.

J. B. Rhine

The founding of the Society for Psychical Research in 1882 was an attempt to organize and professionalize psychical research. Other societies, such as the American Society for Psychical Research quickly followed. Journals and proceedings were published and international congresses were held. Despite these steps towards institutionalization, psychical research continued for the next 50 years to be an uncoordinated activity of amateurs. No agreed upon program or central body of concepts characterized the field.

During this period, psychic researchers disagreed among themselves on issues involving subject matter, methodology, and theory. On one side were those, perhaps the majority, who supported the spiritist hypothesis that psychic phenomena reflected the activity of departed spirits or superintelligent beings. Opposed to these were psychic researchers like Nobel Laureate Charles Richet who defended the position that the phenomena could be explained in terms of a "psychic force" without assuming survival or spirits [47].

Another division was between those who felt that psychical research should confine itself to mental phenomena such as telepathy, premonitions, and clairvoyance. Opposed to these were those who felt that the physical phenomena such as levitation, materialization, poltergeist events, and psychokinesis should be the focus of inquiry. The majority of psychical researchers believed in telepathy but were dubious about clairvoyance. But a strong minority, lead by Richet, believed that clairvoyance not only existed but was the basic phenomenon underlying telepathy.

Possibly the most divisive issue of all was the question of what sort of a research program was appropriate for psychical investigation. A small, but vocal minority wanted psychical research to become a rigorous experimental science. A larger group felt that the natural-historical method was more appropriate because so many of the important phenomena were spontaneous and not observable in the laboratory. Opposed to both these groups were members of the societies who felt that the quantification and rigor of the natural sciences were irrelevant to the study of psychical phenomena.

The event that is credited with providing psychical research with a common focus and a coherent research program was the publication in 1934 of J. B. Rhine's mongraph *Extra-Sensory Perception* [46]. Mauskopf and McVaugh [47] provide an excellent survey of the period from 1915 to 1940, which they treat as the period when psychical research made the transition from a pre-paradigmatic to a paradigmatic research program.

Rhine pulled together the various strands already existing in psychical research and coordinated them into a coherent program. He also coined the terms "parapsychology" to refer to the new experimental science which descended from psychical research and "extra-sensory perception" to refer to the basic phenomenon which was to be studied. In agreement with Richet, and in disagreement with the British parapsychologists, Rhine viewed clairvoyance as on the same footing with telepathy. Later, precognition was also put under the rubric of extra-sensory perception (ESP). ESP became defined as "Knowledge of or response to an external event or influence not apprehended through known sensory channels" [48]. This included telepathy, clairvoyance, precognition, and retrocognition. The psychic phenomena not involving reception of information were included under the term "psychokinesis" (PK) which is defined as "The influence of mind on external objects or processes without the mediation of known physical energies or forces" [48]. Today both ESP and PK are included under the more general term "psi" which is "A general term to identify a person's extrasensorimoter communication with the environment" [48].

Rhine's 1934 mongraph deals only with clairvoyance and telepathy. In 1934 he also began research programs on precognition and psychokinesis. Apparently, he was reluctant to publicize these latter programs too soon for fear of making parapsychology too controversial and unacceptable to mainstream science [48]. He waited until 1938 before he published anything on precognition and until 1943 for the first reports on his PK results.

The major innovation introduced by Rhine was the use of the five target designs: circle, cross, wavy lines, square, and star. These patterns were printed on cards and the standard ESP deck consisted of 5 cards of each symbol for a total of 25 cards. Rhine also introduced standard procedures for using these target materials. The two most common were the Basic Technique and the Down Through Technique. In the Basic Technique (B.T.), the deck is shuffled and placed face down, the percipient guesses the value of the top card; this is then removed and laid aside and the percipient guesses the value of the second card; the second card is then removed and laid on top of the first and the percipient now guesses the third card; etc. This procedure is continued until all 25 cards have been used. At the end of such a "run," a check is made to see how many guesses were hits. If the procedure was supposed to test telepathy then an agent would look at each card at the time the percipient was trying to guess its symbol. If clairvoyance was being tested, no one would look at each card as it was placed aside. The Down Through Technique (D.T.) tested clairvoyance by having the percipient guess the symbols from top to bottom before any of them were removed for checking against the call. The D.T. technique is considered to be superior methodologically in that it better protects against inadvertent sensory cues from the backs of the cards.

Extra-Sensory Perception attracted the attention of both the psychical researchers and the skeptics for two reasons. Rhine's database consisted of 91 174 separate trials or guesses over a three-year period using a number of nonprofessional individuals as percipients. More important was the unprecedented level of success which he reported. Of the 85 724 guesses recorded using the five-symbol ESP decks, 24 364 were "hits." This was 7219 more hits than the 17 145 that would be expected just by chance. The odds against this being just an accident are calculated as being practically infinite. His subjects averaged 7.1 hits per run of 25 as against the chance expectation of 5. Although this is only 2 extra hits per 25, such consistency over this huge number of trials and different subjects had no precedent in the prior history of psychical research.

Rhine's best subject, Hubert Pearce averaged 8 hits per run over a total of 17 250 guesses. As Rhine notes [46]:

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

> Most people are more impressed by a spectacular series of successive hits than by lower but cumulative scoring. Pearce's scoring 25 straight hits under clairvoyant conditions, in my presence, and Zirkle's 26 straight hits in pure telepathy with my assistant, Miss Ownbey, are the best instances of these. Other subjects have approached these. Linzmayer scored 21 in 25 clairvoyance, in my presence; Miss Ownbey herself, unwitnessed, scored 23, pure clairvoyance. Miss Turner's score of 19 in distance P.T. [pure telepathy] work stands out because of the 250 miles between her and the agent. Miss Bailey scored 19 in P.T. in the same room with the agent, as did also Cooper. The odds against getting one series of 25 straight hits by mere chance would be 5^{25} which is nearly 300 quadrillions—just one score of 25! A small part of our 90 000 trials.

Rhine's work provided the model for most parapsychological work from 1934 to around 1970. Using card-guessing with the five ESP symbols, an astonishing variety of questions about ESP were investigated [48]. Because of its huge database, its claims to statistical and experimental sophistication, and its unprecedented rate of success Rhine's research gained the attention of scientific and popular audiences [47]. At first scientists were at a loss about how to react. Many scientists, as a result of reading Rhine's work, were encouraged to try to replicate the results. A few got encouraging results, but most failed.

The first attacks by the critics were aimed at Rhine's statistical procedures. As it turned out, some of Rhine's statistical procedures were technically incorrect, but, for the most part, his results could not be explained away as due to inappropriate statistical procedures. The critics turned out to be wide off the mark in many of their accusations. On the whole, however, the statistical debate led to constructive developments and improved clarification about the proper use of statistical procedures in such experiments [47].

Having essentially lost the statistical battle, the critics then turned to Rhine's experimental controls. Here, he was much more vulnerable. And, ironically, it was the British psychical research community that had anticipated the critics and which provided the sharpest critiques of Rhine's methods [47]. The British parapsychologists were astonished both by Rhine's apparent ease in finding successful percipients as well as his claims that clairvoyance worked as well as telepathy. With only a few exceptions, they had found only evidence for telepathy. And their experience had convinced them that telepathic powers were very rare. While they welcomed Rhine's contribution, they were quick to point out many of its defects, especially Rhine's inadequate description of his procedures and the seeming casualness of his experiments.

During the 1930s, nevertheless, Rhine's work as reported in *Extra-Sensory Perception*, was hailed by parapsychologists as the best scientific case for ESP ever put before the world. Today, as I understand it, most parapsychologists, although they acknowledge its seminal influence on the development of the field, dismiss much of Rhine's earlier work as nonevidential because of its loose controls, poorly made target materials, and inadequate documentation.

S. G. Soal

Rhine's strongest critic among the British parapsychologists was the mathematician S. G. Soal. Just prior to the appearance of Rhine's monograph, Soal had conducted a huge series of card-guessing experiments with only chance results. But the experiments for which Soal became most renowned began as a direct response to Rhine's monograph.

After five years of heroic research, Soal was sure that he had succeeded only in demonstrating the laws of chance. A colleague, however, persuaded him to check for a certain trend in his data. And this resulted in a new series of experiments that for almost 25 years were hailed as the most convincing and fraud-proof demonstration of ESP ever achieved. Because the experiment and results seemed so impressive, some critics, in a way reminiscent of Carpenter's attacks upon Wallace and Crookes and within the spirit of Sidgwick's False Dichotomy, openly accused Soal of fraud on no other basis than that his results were too good. Other critics attacked him on grounds that were irrelevant. As it turns out the critics were right, but for the wrong reasons!

As soon as Soal heard about Rhine's successful American research, he began an ambitious program to replicate Rhine's findings in England. Soal started late in 1934 and continued his experiments for five years. At the end he had accumulated 128 350 guesses for 160 percipients. This is almost 30 percent more guesses than Rhine had accumulated for his 1934 monograph. Soal was sure that he had removed all the flaws and weaknesses that had characterized Rhine's work. Unfortunately, Soal found that this enormous effort yielded "little evidence of a direct kind that the persons tested, whether considered as individuals or in the mass, possessed any faculty for either clairvoyance or telepathy" (quoted in [49]).

Soal reported these results to a stunned parapsychological world in 1940. At the same time another British parapsycholgist, Whately Carington, reported the results of telepathy experiments which seemed to show a "displacement effect." Instead of achieving hits on the target, his subjects seemed to achieve above chance matches when their guesses were matched with either the immediately preceding or the next target in the series. Carington asked Soal to check his data to see whether he, too, might find such a displacement effect [49].

Soal was reluctant to do so. He told Goldney that he thought Carington's request was preposterous and he wasn't going to waste his time going through his huge batch of records. But Carington persisted and Soal finally agreed. Soal found, among the records of his 160 percipients, two who seemed to show Carington's displacement effect. Although this finding was published, presumably Soal realized that such a post hoc finding had to be replicated [49].

Fortunately, one of his two percipients, Basil Shackleton, was available for testing during the years 1941 through 1943. With the collaboration of K. M. Goldney, 40 sittings which yielded a total of 11 378 guesses were obtained with Shackleton during this difficult period when England was at war. As had been the case with the original testing, Shackleton's guesses were at chance level when compared with the actual target, but when compared with the symbol coming up immediately after the target (precognitive hitting), Shackleton's guesses yielded 2890 successes as compared with the 2308 expected by chance. The odds against this being a chance occurrence were calculated to be more than 10^{35} to 1 [50].

In 1945 Soal was able to begin experimenting on the second percipient who had displayed the displacement effect in the original data, Mrs. Gloria Stewart. He was able to accumulate a total of 37 100 guesses during 130 separate sittings. Unlike Shackleton's or her own previous perfor-

 Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

mance, her hitting this time was on the actual target rather than on the immediately preceding or following trial. She managed to achieve 9410 hits which were 1990 more hits than would be expected by chance. The odds against such a result were calculated as 10^{70} to 1 [50].

Soal's stated objective was to make these experiments completely error-free and fraudproof. The basic procedure, which was varied slightly on occasion, was as follows. The percipient—Basil Shackleton or Gloria Stewart—sat in one room monitored by one of the experimenters (EP). In an adjoining room, the sender or agent sat at table opposite the second experimenter (EA). The door between the rooms was slightly open so that the percipient could hear EA's call as to when to make his or her guess. The percipient, of course, could see neither the agent or EA. A screen, with a small aperture separated the agent and EA. For each block of 50 trials EA had before him a list of randomized numbers which determined the target for each trial. Each number could range from 1 to 5. If the target number for the first trial was, say, 3, EA would hold up a card with the number 3 on it so that it could be seen by the agent through the aperture. The agent had lying before him in a row, five cards. Each card had a different drawing of an animal on it: elephant, giraffe, lion, pelican, and zebra. Before each block of trials, the agent shuffled the order of the picture cards. If EA held up a card with 3 on it, the agent would turn up the third card and concentrate upon the animal depicted on it. The percipient would then try to guess which animal was being "sent" and write his guess for that trial in the corresponding place on the response sheet. After every block of 50 trials, the agent reshuffled the target cards so that, for that block, only the agent knew which animal corresponded with which number.

In addition to this rather elaborate arrangement, independent observers were invited to attend many of the sittings. Several professors and a member of parliament were among the observers. On some blocks of trials, unknown to the percipient, the agent did not look at the symbols. This was a test for clairvoyance. Other variations were introduced from time to time. The experiments with Gloria Stewart, while following the same pattern, were admittedly not as carefully controlled. Special precautions were also introduced to ensure that the prepared target sequences could not be known to agent or percipient in advance. And careful safeguards were introduced during the recording of the results and the matching of the targets against the guesses. Duplicates of all records were made and posted immediately after each session to a well-known academic.

Never before had so many safeguards been introduced into an ESP experiment. With so many individuals involved, and with prominent observers freely observing, any form of either unwitting cueing or deliberate trickery would seem to be just about impossible. If fraud of any sort were to be suspected, it would seemingly require, under the stated conditions, the active collusion of several prominent individuals. Beyond these safeguards, Soal randomized his targets, instituted sophisticated checks for randomness, and used the most appropriate statistical procedures. Despite these elaborate precautions, the two subjects managed to consistently score above chance over a number of years.

Soal's findings were hailed as definitive by the parapsychological community and were so good that the rest of the scientific community, including the skeptics, could not ignore them. Here was one of Rhine's severest critics, a man who had spent many years meticulously conducting enormous card-guessing experiments with only chance results, a man who was by profession a mathematician, and an experimenter who had seemingly taken every known precaution to guard against every loophole and possibility of error, who suddenly demonstrated highly successful telepathic and precognitive results over sustained periods of time with two percipients.

Whately Carington, the parapsychologist who convinced Soal to re-examine his seemingly unsuccessful results, wrote (as quoted in [51]):

> Mr. Soal is a most remarkable man, for whose work I have the highest possible admiration. Possessed of a more than Jobian patience, and a conscientiousness, thoroughness which I can only describe as almost pathological, he worked in various branches of the subject for many years with nothing but a succession of null results to show for it... Hoping to repeat Rhine's experiments in England, he tested 160 persons, collecting 128 350 Zener card guesses single-handed, and using the most elaborate precautions against every possible source of error... If I had to choose one single investigation on which to pin my whole faith in the reality of paranormal phenomena, or with which to convince a hardened skeptic (if this be not a contradiction in terms), I should unhesitatingly choose this series of experiments, which is the most cast-iron piece of work I know, as well as having yielded the most remarkable results.

Similar sentiments were expressed by virtually every parapsychologist who commented on this work. As just one illustration, R. A. McConnell [52] phrased it as follows:

> As a report to scientists this is the most important book on parapsychology since the 1940 publication of *Extra-Sensory Perception After Sixty Years*. If scientists will read it carefully, the 'ESP controversy' will be ended.

G. R. Price's Critique

Ironically, some critical scientists did read it carefully but, contrary to McConnell's prognosis, the controversy did not end. Indeed, one of the first major reviews in a scientific journal raised the controversy to new heights. Although the Shackleton experiments had originally been reported by Soal and Goldney in the *Proceedings of the Society for Psychical Research* in 1943, the scientific world did not become aware of those experiments until they were reported along with the later experiments with Gloria Stewart in the 1954 book *Modern Experiments in Telepathy* by Soal and Bateman [50].

What fueled the controversy was an unprecedented review article, nine pages in length, appearing in *Science*, the prestigious journal of the American Association for the Advancement of Science. On August 26, 1955 George R. Price's article on "Science and the Supernatural" was the only feature article for that issue. Price, who as far as I can tell had never before written on parapsychology, was described as being a research associate in the Department of Medicine at the University of Minnesota.

Price began his controversial article by stating that, "Believers in psychic phenomena—such as telepathy, clairvoyance, precognition, and psychokinesis—appear to have won a decisive victory and virtually silenced opposition" [53]. Price writes that such a victory has seemed close in the past, but always critics have managed to find flaws. But

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Price sees the time at which he is writing as unique because practically no scientific papers had attacked parapsychology during the preceding 15 years [53].

> The victory is the result of an impressive amount of careful experimentation and intelligent argumentation. The best of the card-guessing experiments of Rhine and Soal show enormous odds against chance occurrence, while the possibility of sensory clues is often eliminated by placing cards and percipient in separate buildings far apart. Dozens of experimenters have obtained positive results in ESP experiments, and the mathematical procedures have been approved by leading statisticians.
>
> I suspect that most scientists who have studied the work of Rhine (especially as it is presented in *Extra-Sensory Perception After Sixty Years* . . . and Soal (described in *Modern Experiments in Telepathy*), . . . have found it necessary to accept their findings. . . . Against all this evidence, almost the only defense remaining to the skeptical scientist is ignorance, ignorance concerning the work itself and concerning its implications. The typical scientist contents himself with retaining in his memory some criticism that at most applies to a small fraction of the published studies. But these findings (which challenge our very concepts of space and time) are—if valid—of enormous importance, both philosophically and practically, so they ought not to be ignored.

Price then elaborates upon a suggested scheme, using redundancy coding, which would make ESP useful, even if it is a very weak and erratic form of communication. He then presents his version of Hume's argument against miracles. He quotes Tom Paine's more succinct version of the same argument, ". . . is it more probable that nature should go out of her course, or that a man should tell a lie?"

To justify using Hume's argument as his only grounds for accusing the parapsychologists of cheating, Price first tries to show that if ESP were real it would violate a number of fundamental principles underlying all the sciences. Some of these principles are that the cause must precede the effect, signals are attenuated by distance, signals are blocked by appropriate shielding, and so forth. ESP, according to Price, if it exists, violates all these principles. Then Price puts forth reasons why he considers ESP to be a principle of magic rather than merely a previously undiscovered new law of nature. "The essential characteristic of magic is that phenomena occur that can most easily be explained in terms of action by invisible intelligent beings . . . The essence of science is mechanism."

These lengthy considerations back up Price's solution to coping with the challenge of parapsychological claims [53]:

> My opinion concerning the findings of the parapsychologists is that many of them are dependent on clerical and statistical errors and unintentional use of sensory clues, and that all extrachance results not so explicable are dependent on deliberate fraud or mildly abnormal mental conditions.

Actually, nothing is novel or startling about Price's opinion. The same opinion, stated in just about the same words, probably is held by all skeptics. Price has carried his opinion beyond skepticism, however. The thrust of his article is that the *best* research in parapsychology as exemplified in the work of Rhine and Soal *cannot* be dismissed on the basis of "clerical and statistical error and unintentional use of sensory clues." Therefore, he concludes that the results of this otherwise exemplary research *must* be due to fraud. He does not feel that he requires any evidence of fraud. Hume's argument against miracles gives him sufficient license. Price's position, of course, no longer belongs to skepticism but to dogmatism. His position seemingly is that no research, no matter how well done, can convince him of ESP.

But Price does not want to go to quite that extreme. He says that he still can be convinced *provided that* the parapsychologists can supply him with just one successful outcome from a truly fraudproof experiment. "What is needed is one completely convincing experiment— just one experiment that does not have to be accepted simply on the basis of faith in human honesty. We should require evidence of such nature that it would convince us even if we knew that the chief experimenter was a stage conjurer or a confidence man."

But does not the Soal experiment with Shackleton and Stewart meet this criterion? No, says Price, because he can imagine scenarios in which cheating could have taken place. Price then presents a number of possible ways that he feels cheating *could* have occurred in the Soal experiments [53].

> I do not claim that I know how Soal cheated if he did cheat, but if I were myself to attempt to duplicate his results, this is how I would proceed. First of all, I would seek a few collaborators, preferably people with good memories. The more collaborators I had, the easier it would be to perform the experiments, but the greater would be the risk of disclosure. Weighing these two considerations together, I'd want four confederates to imitate the Shackleton experiments. For imitating the Stewart series, I'd probably want three or four—although it is impossible to be certain, because the Stewart sittings have not been reported in much detail. In recruiting, I would appeal not to desire for fame or material gain but to the noblest motives, arguing that much good to humanity could result from a small deception designed to strengthen religious belief.

After providing a sampling of scenarios in which cheating could have occurred, all involving the collusion of three or more investigators, participants and onlookers, Price supplies some designs of what he would consider to be a satisfactory test. The key to all his designs involves a committee. "Let us somewhat arbitrarily think of a committee of 12 and design tests such that the presence of a single honest man on the 'jury' will ensure validity of the test, even if the other 11 members should cooperate in fraud either to prove or disprove occurrence of psi phenomena."

Perhaps if some enterprising group of scientists collaborated and conducted an ESP experiment with positive results according to one of Price's approved designs, the outcome might very well convince *him*. But I do not think it would, nor should it, convince the majority of skeptical scientists. Without going into all its other faults, a single experiment—no matter how elaborate or allegedly fraudproof—is simply a unique event. Scientific evidence is based on cumulative and replicable events across laboratories and investigators. The rubbish heap of scientific history contains many examples of seemingly air-tight experiments whose results have been discarded because later scientists could not replicate the results. The experiments on mitogenetic radiation would be just one example. No one has found fault with the original experiments. But since later experimenters could not replicate the results, the original experiments have been cast aside. Can anyone doubt that this would not also happen to a successful, but nonreplicable, ESP outcome from one of Price's "satisfactory tests?"

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Price tells us, "that I myself believed in ESP about 15 years ago, after reading *Extra-Sensory Perception After Sixty Years*, but I changed my mind when I became acquainted with the argument presented by David Hume in his chapter 'Of miracles' in *An Enquiry Concerning Human Understanding*." So Hume supplies him with his escape hatch.

But all this seems unnecessarily dramatic. Price has fallen into a particularly stark version of the False Dichotomy. He has been forced into the very position that Henry Sidgwick wanted for the critics. The best ESP evidence is so good that either the critic must admit the reality of psi or accuse the proponents of lying and fraud. In falling into this trap, one that critics from the days of Hare and Crookes right up to the present keep falling into, Price has needlessly attributed to the Rhine and Soal results a level of evidential value which they cannot carry. At the same time, Price has implied that he is sufficiently expert in parapsychological research that he can infallibly judge when a given outcome unquestionably supports the conclusions of the experimenters. In fact, I doubt that even the parapsychologists are ready to give such power to a single experiment, even one so seemingly well-conducted as Soal's.

Price writes as if, when confronted with experimental evidence for psi, such as can be obtained by reading *Extra-Sensory Perception After Sixty Years* or *Modern Experiments in Telepathy*, he must immediately a) find ways to reject the findings on the basis of possible sensory leakage, statistical artifacts, or loose experimental controls; or b) accept the outcome as proof of psi; or c) accuse the investigators of fraud if he can imagine some scenario, no matter how complex and unlikely, under which fraud *could* have occurred. Price just does not understand either parapsychological research or scientific research in general if he truly believes these are the only alternatives open to him. Unfortunately, Price is behaving like many of the other outspoken critics of psychical research. To Price's credit, he has at least tried to make his basis for action explicit.

Both Rhine and Soal, in their responses to Price's critique, eagerly accepted Price's implicit endorsement of their experimental procedures. Soal commented that, "It is very significant and somewhat comforting to learn that Price admits that 'most of Soal's work' cannot be accounted for by any combination of statistical artifact and sensory leakage" [54]. Soal also examined in detail Price's various proposed schemes for faking the experiments [54]:

> Price goes to great length in devising variations on this theme, but they all depend on the Agent being in collusion with the chief Experimenter or with the Percipient. Now four of the Agents with whom Mrs. Stewart was highly successful were lecturers of high academic standing at Queen Mary College in the University of London. Two were senior lecturers and the other two were mathematicians who had done distinguished creative work. A fifth Agent who was brilliantly successful over a long period was a senior civil servant, in fact an assistant director of mathematical examinations in the Civil Service. Now is it plausible to suppose that I, as chief Experimenter, could persuade any of these men to enter into a stupid and pointless collusion to fake the experiments over a period of years? What had any of them to gain from such deplorable conduct? If I had gone to any of them and suggested (as Price recommends) that in a good cause a little deception would do no harm, I know quite plainly that the result would have been a first-class scandal in university circles.

Rhine found even more solace in Price's attack. "Strange though it may seem, the publication of the George Price paper . . . is, on the whole, a good event for parapsychology" [55]. For one thing it was a way of getting a lot of instruction on parapsychology before the scientific community. Rhine also felt Price's vivid portrayal of the potential importance of ESP was valuable. He welcomed Price's effective rebuttal against the standard criticisms against ESP. And Rhine especially liked the fact that Price focussed on the point that psi was incompatible with the materialism of science [55]:

> [Price] even more than any other critical reviewer gives indication of having felt the force of the evidence for ESP. When he turns then—albeit a bit too emotionally—and says that, according to the current concept of nature, ESP is impossible and therefore the parapsychologists must all be fakers, he at least draws the issue where it can be squarely met. The answer of the parapsychologist is: "Yes, either the present mechanistic theory of man *is* wrong—that is, fundamentally incomplete—or, of course, the parapsychologists *are* all utterly mistaken." *One* of these opponents is wrong, take it, now, from the pages of *Science*! This recognition of the issue gives point to the findings of parapsychology in a way none can easily miss.

Notice that Rhine and Price agree on some aspects of this controversy. Both Rhine and Price believe that if the claims of parapsychology are correct the foundations of science are seriously threatened. Rhine welcomes such a destruction of what he calls materialism. Price seems willing to take the most drastic measures to avoid this overthrow of what he calls the basic limiting principles. (Not all parapsychologists agree with Rhine that the acceptance of psi need be inconsistent with scientific materialism.) One issue involves what it means for contemporary science to accept the reality of psi. This concerns matters that are currently controversial among philosophers of science. And so, it is probably not fruitful to attempt to deal with them here.

Rhine and Price also agree that the standard arguments against parapsychological evidence do not hold up. According to reasonable scientific criteria, the evidence for psi is more than adequate. And so it is at this point that both Rhine and Price want to have the showdown. Price, as a defender of the materialistic faith, puts all his money on the hope that the parapsychologists have faked the data. He has no evidence to back this claim. But if he can invent possible scenarios wherebye trickery *might* have been committed in a given experiment, then he believes he can, under license from David Hume, assume that fraud must have taken place. He is not completely dogmatic about this. If the parapsychologist can come up with positive results in at least one experiment conducted under what Price considers to be fraudproof conditions, then Price has committed himself to accept the consequences.

Many issues are raised by Price's dramatic confrontational posturing. At this point, I will just mention one. Price goes beyond conventional scientific practice when he empowers a given experiment with the ability to prove the existence of psi. Once we realize that no experiment by itself definitely establishes or disproves a scientific claim, then Price's extreme remedies to save his image of science become unnecessary. No matter how well-designed and seemingly flawless a given experiment, there is always the possibility that future considerations will reveal previously unforseen loopholes and weaknesses.

Indeed, a careful analysis of the Soal experiment will reveal a variety of weaknesses. For example, in spite of the number of observers and experimenters, Soal always had control over the prepared target sequences or over the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

basic recording. And both Shackleton and Stewart only produced successful results when Soal was present. On one occasion, without informing Soal, his co-investigator Mrs. Goldney conducted a sitting with Shackleton. The outcome was unsuccessful. The American parapsychologist J. G. Pratt ran a series of experiments with Mrs. Stewart without Soal's presence. No evidence for psi was found. And whereas all Rhine's results showed no difference between telepathic and clairvoyance trials, both Shackleton and Mrs. Stewart produced successful results only on telepathic trials. Furthermore, in spite of the much vaunted measures to guard against sensory leakage, the actual experimental setup, when carefully considered, offered a variety of possibilities for just such unwitting communication.

None of the foregoing considerations, in themselves, account for Soal's findings. But they make superflous, I would argue, the hasty assumption that the findings can only be explained either by psi or some elaborate form of dishonest collusion.

The Discrediting of Soal

As it turns out, if Soal did cheat—and it now seems almost certain that he did, he almost certainly did so in ways not envisaged by either Price or Hansel. The scenarios generated by these two critics involved collusion among several of the principals. Soal apparently managed the fraud entirely on his own, or, at most, with the collusion of one other person. Furthermore, he probably used a variety of different ways to accomplish his goals.

If it had not been for a series of seemingly fortuitous events, Soal's experiment might still occupy the honored place in the parapsychologists' exhibits of evidence for psi [56]–[60]. The discrediting of Soal's data occurred through a number of revelations during the period from 1955 through 1978. Up until 1978 the accumulation of evidence suggested that something was highly suspicious about the records in the Shakleton experiments. The case was strong enough to discredit Soals' results in the judgment of some leading parapsychologists, but many others still defended Soal's findings.

The final blow to the credibility of Soal's results came in 1978 when Betty Markwick published her article "The Soal-Goldney experiments with Basil Shackleton: New evidence of data manipulation" [60]. As with the previous revelations of peculiarities in the data, Markwick's stunning findings arose out of a series of fortuitious incidents.

The story is much too complicated to relate here. Essentially, Markwick had begun a rather elaborate project to clear Soal of the accumulating charges that he had tampered with the data. Her plan involved searching the records with the aid of a computer to find subtle patterns which, if they existed, would account for the anomalies found by the critics and would vindicate Soal. Markwick did not find such patterns. Instead, she discovered previously unnoticed patterns that could be accounted for if one assumed that Soal had used a sophisticated plan for inserting "hits" into the records while he was apparently summarizing and checking the results. Reluctantly, she was forced to conclude that only the hypothesis of deliberate tampering with the data could explain her findings [60].

> Protestations to the effect that Soal, a respected scientist, would not have cheated in his own experiments—and that anyway the rigorous experimental conditions in the Shackleton series precluded fraud—seem to me to carry little weight in the face of the evidence. We can rarely [illegible] and perhaps Soal was as clever. It is futile to argue that the prison cell is escape-proof when the inmate has clearly gone.

Markwick, obviously dismayed at having discovered that Soal almost certainly faked his data suggests two possible explanations for why he might have done so. One of her hypotheses made use of the well-known fact that Soal sometimes did automatic writing in a dissociated state. Markwick suggested the possibility that Soal may have had a split personality and that the cheating was done by his other self.

Markwick's second hypothesis involved data massage and has more universal psychological plausibility (although it is not necessarily inconsistent with her first hypothesis). She assumes that Soal's enormous accumulation of negative ESP findings were obtained legitimately. She also assumes that his *post hoc* finding of consistent displacement effects in the data of Basil Shackleton and Gloria Stewart was also legitimate [60].

> Having embarked upon the Shackleton series, one may imagine the scoring rate begins to fade (as ESP scores are wont to do after the initial flush of success). Soal, seeing the chance slipping away of gaining scientific recognition for Parapsychology, a cause in which he passionately believes, succumbs to the temptation of "rectifying" a "temporary" deficiency.

Markwick's second scenario is consistent with known patterns in which scientists have tampered with their data [61], [62]. The components appear to be: 1) the investigator believes, on the basis of previous experience, that the phenomenon under investigation is "real"; 2) for some unknown reason his current research fails to reveal the phenomenon; 3) if he reports negative results his readers might wrongly believe that the phenomenon does not exist; 4) as a result, the "truth" and assumed positive consequences of the phenomenon might be lost to humanity. Given these ingredients, it takes a very small step for the investigator to convince himself that he is helping both the truth and a good cause along by doctoring his data.

William James, with reference to his experiences in psychical research, suggested that cheating in order to convince others of the "reality" you know to be the case might be defensible. James discussed this matter in his last essay on psychical research. He referred to the policy of English investigators to consider a medium who has been caught cheating as one who always cheats. He indicated that he thought this had generally been a wise policy [2].

> But, however wise as a policy the S.P.R.'s maxim may have been, as a test of truth I believe it to be almost irrelevant. In most things human the accusation of deliberate fraud and falsehood is grossly superficial. Man's character is too sophistically mixed for the alternative of "honest or dishonest" to be a sharp one. Scientific men themselves will cheat—at public lectures—rather than let experiments obey their well-known tendency towards failure.

James gave two examples of such cheating. And then revealed the following about his own behavior [2]:

> To compare small men with great, I have myself cheated shamelessly. In the early days of the Sanders Theater at Harvard, I once had charge of a heart on the physiology of which Professor Newell Martin was giving a popular lecture. This heart, which belonged to a turtle, supported an index-straw which threw a moving shadow, greatly enlarged, upon the screen, while the heart pulsated. When certain nerves were stimulated, the lecturer said, the heart would act in

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

certain ways which have been found to be ... far gone and, although it stopped duly when the nerve of arrest was excited, that was the final end of its life's tether. Presiding over the performance, I was terrifed at the fiasco, and found myself suddenly acting like one of those military geniuses who on the field of battle convert disaster into victory. There was no time for deliberation; so, with my forefinger under a part of the straw that cast no shadow, I found myself impulsively and automatically imitating the rhythmical movements which my colleague had prophesied the heart would undergo. I kept the experiment from failing; and not only saved my colleague (and the turtle) from humiliation that but for my presence of mind would have been their lot, but I established in the audience the true view of the subject. The lecturer was stating this; and the misconduct of one half-dead specimen of heart ought not to destroy the impression of his words. "There is no worse lie than a truth misunderstood," is a maxim which I have heard ascribed to a former venerated President of Harvard. The heart's failure would have been misunderstood by the audience and given the lie to the lecturer. It was hard enough to make them understand the subject anyhow; so that even now as I write in cool blood I am tempted to think that I acted quite correctly. I was acting for the *larger* truth, at any rate, however automatically... To this day the memory of that critical emergency has made me feel charitable towards all mediums who make phenomena come in one way when they won't come easily in another. On the principles of the S.P.R., my conduct on that one occasion ought to discredit everything I ever do, everything, for example, I may write in this article—a manifestly unjust conclusion.

I wonder if James would have approved of the way William Crookes covered up the cheating of the medium Mary Showers in behalf of "the larger truth?" Mary Showers, a young medium, conducted at least one joint seance with Florence Cook in Crookes' home. Apparently Crookes had several other sittings with Mary. Daniel Home presumably heard rumors that Crookes might be having an affair with the young Mary Showers. Crookes wrote a letter to Home explaining how the scandal had originated [63].

> According to Crookes he had obtained a complete confession from Mary Showers in her own handwriting that her phenomena were wholly dependent upon trickery and the ocassional use of an accomplice. Crookes said, however, that he had undertaken not to reveal the fact that Mary was fraudulent even to her own mother, because of "the very great injury which the cause of truth would suffer if so impudent a fraud were to be publicly exposed."

The Post-Rhine Era

Rhine's card-guessing paradigm dominated experimental parapsychology from 1934 to at least the 1960s. Since the 1960s card-guessing experiments have played a minor role. Contemporary parapsychologists have deviated from Rhine's paradigm in a variety of ways. In Rhine's paradigm both the possible targets and the possible responses are severely restricted. The targets consist of five, deliberately neutral and simple, symbols. And, on each trial, the percipient is restricted to calling out the name of one of these possible five symbols. From a strictly methodological viewpoint these restrictions have several advantages. Most percipients have no strong preferences for any of the symbols; randomizing of targets is straightforward; scoring of hits and misses is unambiguous; and the statistical calculations are fairly standard.

But these same features have been blamed by contemporary investigators for the lack of impressive findings since the spectacular scoring reported by Rhine in 1934 [46]. Because the symbols are relatively meaningless and uninteresting, the repetitive guessing over many trials is boring and, according to the parapsychologists, contributes to both a lack of motivation and emotional involvement which might be needed for the effective functioning of psi.

As a result, one break with the past is the increased use of more complex and meaningful targets such as reproductions of paintings, travel slides, geographical locations, and emotionally laden photographs. In addition, instead of the forced-choice procedure of the card-guessing, most experimenters allow free-responding on the part of their percipients. Percipients are encouraged, on a given trial, to free-associate and describe, both in words and in drawings, whatever comes to mind. The use of free responses complicates enormously the problems of scoring and statistical analysis. But parapsychologists believe the added complications are a small price to pay if the newer procedures produce better psychic functioning.

Along with the free-response designs, parapsychologists have renewed their interest in the possibility that psychic functioning may be enhanced in altered states such as dreaming, hypnosis, meditation, sensory-deprived states, and progressive relaxation. The basic idea is that these altered states greatly reduce or block attention to external sensory information while, at the same time, increasing attention to internal mentation. Under such conditions it is hypothesized that the psi signal is easier for the percipient to detect because it has less competition from sensory inputs [64]. One survey of 87 experiments in which percipients were in an altered state found that 56 percent reported significant hitting of targets [65].

Another departure from the Rhine paradigm was stimulated by developments in electronic technology. Psi experiments employing Random Event Generators began in the 1970s. Electronic equipment could be used to generate random targets as well as automatically record the percipient's responses and keep running tallies of the hits. Although such equipment has been used to test ESP, the most widespread use has been in the study of psychokinesis. In such experiments an operator or "psychic" attempts to bias the output of a random event generator by mental means alone. In 1980, May, Humphrey, and Hubbard found reports of 214 such experiments, "74 of which show statistical evidence for an anomalous perturbation—a factor of nearly seven times chance expectation" [66].

A third major departure has been the so-called "Remote Viewing" paradigm [22], [24], [28], [67], [68]. The claims made for the ability of this procedure to consistently demonstrate ESP with a variety of percipients are perhaps the strongest ever put forth by parapsychologists [28].

> Our laboratory experiments suggest to us that anyone who feels comfortable with the idea of having paranormal ability can have it.... In our experiments, we have never found anyone who could not learn to perceive scenes, including buildings, roads, and people, even those at great distances and blocked from ordinary perception.... We have, as of this writing, carried out successful remote viewing experiments with about twenty participants, almost all of whom came to us without any prior experience, and in some cases, with little interest in psychic functioning. So far, we cannot identify a single individual who has not succeeded in a remote viewing task to his own satisfaction.

In a more recent assessement of remote viewing, Targ and Harary assert, "In laboratories across this country, and

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

in many other nations [illegible] have investigated remote viewing. Twenty-three of these investigations have reported successful results and produced statistically significant data, where three would be expected" [68].

A fourth emphasis has been the study of personality correlates of the alleged psi ability [48].

In addition to the experimental programs on altered states, random event generators, remote viewing, and personality correlates contemporary parapsychologists have been actively doing research in other areas. The various chapters in the *Handbook of Parapsychology* provide a good idea of the range of topics [48]. The research on reincarnation, survival after death, paranormal photography, psychic metal bending, poltergeist phenomena, hauntings, and faith healing, while admittedly colorful, does not deserve the serious attention of scientists—at least not in its current state. I suspect that most serious parapsychologists would also not want to rest their case on such research.

Today the parapsychologists who want the scientific establishment to take their work seriously do not offer for inspection the evidence that previous generations of psychic researchers believed was sufficient—the findings of Hare, Wallace, Crookes, Gurney, Rhine, or Soal. Nor do they offer up the reports on reincaration, psychic healing, pararormal photography, spoon bending, psychic detection, and the related phenomena which so readily appeals to the media and the public. Instead, they ask us to look at the trends and patterns which they find in research programs carried out in a variety of different parapsychological laboratories.

Two aspects of this new type of claim are worth noting. One is the admission that a single investigation, no matter how seemingly rigorous and fraud-proof, cannot be acceptable as scientific evidence. The idea of a single "critical experiment" is a myth. The second, and related, aspect is that replicability is now accepted as the critical requirement for admission into the scientific marketplace.

Both proponents and critics have previously assumed, either tacitly or explicitly, that the outcome of a single investigation could be critical. Sidgwick believed that the results of the investigation of the Creery sisters were of this nature. The evidence was so strong, he argued, that the critics either had to now either accept the reality of telepathy or accuse the investigators of fraud [30]. Carpenter, rather than withhold judgment until independent investigators had either succeeded or failed in attempts to replicate Crookes' experiments with Home, acted as if he either had to agree to Crookes' claim or prove that Crookes had been duped. Both Price and Hansel insisted that it would be sufficient for Rhine and Soal to convince them of ESP if a parapsychology could perform successfully a single "fraud-proof" experiment.

The myth of the single, crucial experiment has resulted in needless controversy and has contributed to the False Dichotomy. Flew is just one who has argued convincingly that a single, unreplicated event which allegedly attests to a miracle, is simply a historical oddity which cannot be part of a scientific argument [3].

Apparently not all parapsychologists are convinced that the achievement of a repeatable psi experiment is either necessary or desirable for the advancement of parapsychology. The late J. G. Pratt argued that, "Psi is a spontaneous [illegible] predict precisely when it is going to occur in our carefully planned and rigorously controlled experiments than we can in everyday life psychic experiences. . . . Predictable repeatability is unattainable because of the nature of the phenomena" [69].

Pratt argued that parapsychology should give up the quest for the replicable experiment—an impossible goal in his opinion—and concentrate upon accumulating enough data on anomalous happenings to convince scientists and the public that psi is real. Other parapsychologists, however, realize that scientists are not going to be convinced until some semblance of replicability has been achieved. The late Gardner Murphy, while noting that replicability was not necessary for scientific acceptability in some areas of science, argued that for supporting claims for such irrational phenomena as psi, replicability was necessary. And, speaking as one of the dominant figures in parapsychology in 1971, he made it clear that he felt that parapsychology had a long way to go before it achieved replicable results [70].

Perhaps Honorton's position represents the contemporary position of the major parapsychologists [71]:

> Parapsychology will stand or fall on its ability to demonstrate replicable and conceptually meaningful findings. Future critics who are interested in the resolution rather than the perpetuation of the psi controversy are advised to focus their attention on systematic lines of research which are capable of producing such findings.

Psi and Repeatability

As the preceding quotation indicates, Honorton believes that critics should focus on "systematic lines of research" which apparently display replicable and/or "conceptually meaningful" findings. And, as we have seen, contemporary parapsychologists have offered us a number of such systematic lines to demonstrate that they have, in fact, already achieved the goals of repeatability and conceptual meaningfulness. The claims put forth in behalf of the altered state, random event generator, and remote viewing paradigms have already been cited. Similar claims have been made for work on correlates of psi such as attitudes and personality [72].

What can we expect if a critic, in an effort to be open-minded and responsible, accepts the challenge of Honorton and his fellow parapsychologists to examine the accumulated evidence from one or more of the "systematic lines" of inquiry? This challenge opens up a variety of possibilities. Which experiments should be included in the evaluation? It is impractical to consider all the experiments in parapsychology because even in this relatively sparsley populated area the number is by now enormous. In just considering a subset of experiments in the ESP area, Palmer, for example, covered approximately 700 experimental reports [72]. Including PK as well as ESP, I would estimate that, today, a determined critic, who wants to evaluate exhaustively all available experimental reports, might have to cope with upwards of 3000 experiments. Given my recent experience in trying to do justice to just 42 experiments on the Ganzfeld psi phenomenon [73], I would estimate that it could take a responsible critic over five years of almost full-time effort to properly evaluate this material.

Another problem facing both the proponent and critic is,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

met standards of scientific acceptability, but rather, he was assigning flaws if a given deficiency by itself was sufficient to have accounted for the results. And, finally, Akers did not consider the possibility that *combinations* of deficiencies, each in themselves being insufficient, might have been more than enough to account for the reported findings.

Hyman's Critique of the Ganzfeld Experiments

Although Akers' and my critiques were conducted independently, and although our samples and procedures differed in many important ways, we came to essentially the same conclusion. In spite of claims for both scientific confirmation of psi and repeatibility within certain systematic lines of research, both Akers and I concluded that the best contemporary research in parapsychology does not survive serious and careful scientific scrutiny. Parapsychology is not yet ready to bring its case before the general scientific public.

My approach was to look for a research program in parapsychology that consisted of a series of experiments by a variety of investigators and that was considered by parapsychologists as especially promising. I quickly discovered a systematic body of research which many of the leading parapsychologists considered to be the most promising one on the contemporary scene. This research program was based on the Ganzfeld/psi paradigm.

The word "Ganzfeld" is German for total field. It is used to describe a technique in the study of perception which creates a visual field with no inhomogeneities. The motivation for creating such a visual field stems from certain theoretical predictions of Gestalt psychology. A recently developed and simple procedure for creating such a Ganzfeld is to tape halves of ping pong balls over the eyes of subjects. A bright light is then directed to the covered eyes. The percipient experiences a visual field with no discontinuities and describes the perceptual effect as like being in a fog.

The parapsychologists became interested in the Ganzfeld when it was reported that subjects who experience the Ganzfeld quickly enter into a pleasant, altered state. They adopted it as a quick and easy way to place percipients into a state that they felt would be conducive to the reception of psi signals. In a typical Ganzfeld/psi experiment, the percipient has the pin pong balls taped over his eyes and then is placed in a comfortable chair or reclines on a bed. In addition to a bright light shining on the halved ping pong balls, white noise or the sound of ocean surf is fed into the percipient's ears through earphones.

After 15 min or so in this situation, the percipient is presumed ready to receive the psi signal. An agent, in another room or building, is given a target which is randomly selected from a small pool, say, of four pictures (the pool of pictures has been selected, in turn, by random means from a large collection of such pools). The agent concentrates or studies the target during a predetermined time interval. At the same time the percipient, isolated in a relatively sound-proofed chamber, freely describes all the associations and impressions that occur to him during the sending interval.

At the end of the session the halved ping pong balls are removed. The pool of pictures for that trial, including the target, are brought to the percipient. The percipient then indicates on a ranking form how close each of the items in the pool are to the impressions that occurred to him or her during the Ganzfeld session. The most typical scoring procedure classifies the outcome as a "hit" if the percipient correctly judges the actual target as closest to the Ganzfeld impressions.

In the typical experiment a pool of four target candidates is used on each trial. Over a number of trials, the percipients would be expected to achieve hits on 25 percent of the trials just by chance. If the actual rate of hitting is significantly above this chance level, then it is assumed, given that proper experimental controls have been employed, that ESP has probably operated.

Charles Honorton, the parapsychologist who first published a Ganzfeld/psi experiment [76] and who also has strongly defended the paradigm as "psi conducive," responded to my request for cooperation by undertaking to supply me with copies of every relevant report between 1974—the date of the first published Ganzfeld/psi experiment—and the end of 1981—the year I made the request. In January 1982 I received a package containing 600 pages of reports on the Ganzfeld/psi experiment.

The experiments in the database given to me for examination were extracted from 34 separate reports written or published from 1974 through 1981. By Honorton's count, these 34 reports described 42 separate experiments. Of these, he classified 23 as having achieved overall significance on the primary measure of psi at the 0.05 level. This successful replication rate of 55 percent is consistent with earlier estimates of success for this paradigm which ranged from 50 to 58 percent [73]. Approximately half of these experiments had been published in refereed journals or monographs. The remainder had appeared only as abstracts or papers delivered at meetings of the Parapsychological Association. The studies had been authored by 47 different investigators, many of them prominent members of the Parpsychological Association.

The details of my analysis and my conclusions have been published in the *Journal of Parapsychology* [73]. The same issue of that journal contains Honorton's detailed rebuttal to my critique [77]. Here I will merely supply the barebones of my critique.

1) I first examined the claim that the proportion of successful replications of the Ganzfeld/psi experiment was 55 percent. This estimate, it turned out, was based upon a number of questionable assumptions. Much ambiguity exists as to what the unit of analysis should be. In some cases, the individual experimental conditions within a single complicated experiment were each counted as separate "experiments." In other cases, the pooled data over a number of separate experimental conditions were counted as a single unit. That this can make a difference is shown by the fact that when I tried to apply a consistent criterion to the database for determining individual units, I came up with a success rate closer to 30 than to 50 percent. Other considerations such as unknown experiments lead me to conclude that the actual success rate, defining "success" according to Honorton's criterion, was probably around 30 percent.

2) But even a success rate of 30 percent is impressive if the actual rate of success to be expected by chance was the assumed 5 percent. I pointed to a variety of examples in which multiple tests were applied to the same data in such

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

a way as to inflate the actual probability for success just by chance over the a[illegible] Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4 number of factors, I estimated that the actual chance level could easily be 25 percent or higher.

3) In addition to analyses that inadvertently inflated the significance levels, I noted a number of other departures from optimal experimental procedure that could have artificially contributed to the outcomes. These flaws could be clustered into three categories: Security, Statistical, and Procedural. Security flaws included failure to preclude sensory cues as well as loose monitoring of critical aspects in the experiment. Statistical flaws consisted of wrong use of statistical procedures. Procedural flaws consisted of inadequate randomization of targets, incomplete documentation, and possible problems at feedback. What was both surprising and dismaying to me was that not a single experiment in the database was free from at least one of these defects. These defects were chosen to be those that I assume most parapsychologists would agree should not be part of a well-conducted experiment.

4) I tried to make it clear that I was not assuming that these flaws were the cause of the observed results. Rather, I assumed that the presence of such defects could be taken as a symptom that the experiment had not been conducted with adequate care. Indeed, it was clear that at least some of the experiments in the database had been intended to serve only as pilot or preliminary experiments. Nevertheless, I did look at the correlation between the three clusters and success of the experiment. Although the Security and the Statistical clusters did not correlate with outcome, the Procedural cluster did correlate with the probability of obtaining a significant outcome. Honorton strongly disagrees with this conclusion [77].

As a result of my detailed examination of the claims for the Ganzfeld/psi findings, I concluded my long report as follows [73]:

> In conclusion, the current data base has too many problems to be seriously put before outsiders as evidence for psi. The types of problems exhibited by this data base, however, suggest interesting challenges for the parapsychological community. I would hope that both parapsychologists and critics would wish to have parapsychological experiments conducted according to highest standards possible. If one goal is to convince the rest of the scientific community that the parapsychologists can produce data of the highest quality, then it would be a terrible mistake to employ the current Ganzfeld/psi data base for this purpose. Perhaps the Parapsychological Association can lead the way by setting down guidelines as to what should constitute an adequate confirmatory experiment. And, then, when a sufficient number of studies have accumulated which meet these guidelines, they can be presented to the rest of the scientific community as an example of what parapsychology, at its best, can achieve. If studies carried out according to these guidelines also continue to yield results suggestive of psi, then the outside scientific community should be obliged to take notice.

Honorton, not surprisingly, disagrees with my conclusions [77]. After my critique was completed, Honorton carried out a revised and different analysis of the database. He claims his new analysis eliminates my criticisms about inflated significance levels. Honorton also developed his own scale for evaluating the methodological quality of each experiment. According to his ratings, there is no correlation between quality of the experiment and its outcome.

The problem that both of us face when judging the [illegible] is that we are doing this after the fact. Although we agree on several of our ratings, we tend to disagree in ways which suggest our presumed biases. Honorton tends to find more defects in the unsuccessful experiments than I do. On the other hand, I tend to find more defects in the successful experiments than Honorton does. In the absence of double-blind ratings, this aspect of our disagreement represents a stalemate.

However, whether one uses Honorton's or my ratings, the number of departures from accepted methodological procedure is unacceptably high for this database. Although Honorton and I disagree on whether the observed flaws weaken the case for psi, we do not disagree that they exist. So far as I can tell, no parapsychologist has provided an explanation of why almost all of the experiments in this database have at least one of these flaws.

Conclusions

With the exception of the contemporary parapsychological literature, the evidence for psi reviewed in this paper comes from investigations which today's parapsychologists would not put before us as part of their strongest case for psi. Many of these parapsychologists might believe I was being unfair in dwelling upon these castoffs from the past. But it is just this fact that the cases I have examined are now castoffs which brings up important questions about how to approach the contemporary evidence.

Each of the cases from the past which I have discussed were, in their own time, considered to be by the parapsychologists of that day examples of scientifically sound evidence for psi. It is only subsequent generations, for the most part, who have set the preceding exemplars aside. In some cases the reasons for the abandonment of what was once a foundation stone in the case for psi are clear. Subsequent investigators or critics found previously unrecognized defects in the studies or strong suspicions of fraud had been generated. Other experimental paradigms have disappeared from the database for less obvious reasons.

Some previously successful paradigms have disappeared because they no longer seem to yield significant results. Others such as the sheep–goats design seem to have simply gone out of fashion. One major parapsychologist once told me that it seems to be the ultimate fate of every successful paradigm to eventually lose its ability to yield significant results. He believed this was related to the fact that psi depends both upon the novelty of the design and the motivations of the experimenter. At first a new paradigm generates excitement and optimism. But after it has been around for a while, the initial excitement and enthusiasm abates and the experimenter no longer communicates the original emotions that accompanied the paradigm when it was still relatively new.

But, whatever the reason, each generation's best case for psi is cast aside by subsequent generations of parapsychologists and are replaced with newer, more up-to-date best cases. Not only does the evidence for psi lack replicability, but, unlike the evidence from other sciences, it is noncumulative. It is as if each new generation wipes the slate clean and begins all over again. Consequently, the eviden-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

tial database for psi is always shifting. Earlier cases are dropped and replaced with newer and seemingly more promising lines of research. [One of the readers of this paper argues that it is only partially true that parapsychological research is noncumulative. Although his argument might have some validity, I do not think it changes the point I am making here.]

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The late J. G. Pratt, in challenging his parapsychological colleagues' hopes for a repeatable experiment, wrote [69]:

> One could almost pick a date at random since 1882 and find in the literature that someone somewhere had recently obtained results described in terms implying that others should be able to confirm the findings. Among those persons or groups reflecting such enthusiasm are the S.P.R. Committee on Thought-Transference; Richard Hogson (in his investigation of Mrs. Piper); Feilding, Baggally, and Carrington (in their Palladino investigations); J. B. Rhine (work reported in *Extra-Sensory Perception*); Whately Carington (in his work on paranormal cognition of drawings); Gertrude Schmeidler (in her sheep-goat work); Van Bussbach, and Anderson and White (in their research on teacher-pupil attitudes); the Maimonedes dream studies; the Stepanek investigators; the investigators of Kulagina's directly-observable PK effects; research using the ganzfeld technique; and the SRI investigators ("remote viewing"). One after another, however, the specific ways of working used in these initially successful psi projects have fallen out of favor and faded from the research scene—except for the latest investigations which, one may reasonably suppose, have not yet had enough time to falter and fade away as others before them have done.

When Pratt wrote those words in 1978, the "latest investigations" included the Ganzfeld/psi experiments, the Remote Viewing investigations, and the PK research using Random Event Generators. These would have been among the contemporary investigations which, given Pratt's pessimistic extrapolations, "one may reasonably suppose, have not yet had enough time to falter and fade away as others before them have done." Today, signs do seem to indicate that these seemingly "successful" lines of research may be much weaker than had been previously advertised [24], [74], [75].

However, as always, new and more promising lines of work seem to be ready to take their place. Honorton and his colleagues at the Psychophysical Research Laboratories in Princeton, NJ, seem to be developing a number of very promising lines of research [78]. They have been developing a completely automated version of the Ganzfeld experiment which eliminates many of the problems raised by my critique. They have also been perfecting a "transportable" experiment—one that can be carried out by any investigator who has access to an Apple personal computer. The experiment, also completely automated, is a variation of the Random Event Generator paradigm but with a variety of built-in safeguards which apparently eliminate almost all the options for multiple testing.

Nearby, but completely independent of the work going on at the Psychophysical Laboratories, is the research on anomalous phenomena being carried out by Robert Jahn and his associates in the School of Engineering and Applied Science at Princeton University [1], [79], [80]. For more than five years Jahn and his associates have been perfecting the instrumentation and experimental designs for conducting sophisticated variations of both the remote viewing paradigm and the PK work with random event generators. Although they have collected large databases for each of these paradigms, most of the work has been reported only in technical reports. The reported findings do seem impressive, but they have yet to be described in sufficient detail for a full-scale evaluation. And, given both the scale of the effort and the sophistication of the methodology and instrumentation, it will be many years before adequate replications in independent laboratories will be possible.

As promising as this most recent work by Honorton and Jahn might seem to be, none of it has reached a stage where it is ready for a full-scale critical evaluation. Already, the sharp-eyed critic can detect both inconsistencies with previous findings in the same lines of research and departures from ideal practice. As the history of parapsychology teaches us, we will have to wait for several more years before we can adequately judge if somehow these latest efforts can avoid the fate that all their promising predecessors have suffered.

Perhaps, however, history does not have to repeat itself in all its depressing aspects. And I can see some encouraging signs of breaks with previous patterns in the way proponents carry out and defend their findings and the way critics respond.

Since its inception as an institutionalized undertaking, psychical research has suffered from the lack of relevant, informed, and constructive criticism. This particular deficiency seems to be changing. For one thing, the younger generation of parapsychologists have produced some internal critics who are both knowledgeable and effective. In addition to Akers, there are others such as Susan Blackmore, Adrian Parker, Gerd Hövelmann, and J. E. Kennedy who have recognized the current deficiencies of parapsychological research and have a strong committment to raising the standards. Although it is still difficult to find external critics who are both informed and constructive, one can see some indications that this situation may also improve.

Another positive sign is the attempt to replace subjective, impressionistic evaluations of the parapsychological literature with more systematic, explicit assessments. Both Honorton [77] and I [73] have used "meta-analysis" in our dispute over the adequacy of the Ganzfeld/psi database. "Meta-analysis" is a term coined to describe the approach to reviewing a body of research which makes the various phases as explicit and quantitative as feasible [81], [82].

> The approach to research integration referred to as "meta-analysis" is nothing more than the attitude of data analysis applied to quantitative summaries of individual experiments. By recording the properties of studies and their findings in quantitative terms, the meta-analysis of research invites one who would integrate numerous and diverse findings to apply the full power of statistical methods to the task. Thus it is not a technique; rather it is a perspective that uses many techniques of measurement and statistical analysis.

(From [81].)

Meta-analysis is by no means a panacea. Much subjectivity remains on such matters as which studies to include and exclude from the sample, how to score the "effect size" or degree of success of a study, what variables to include, how to assign studies values on the variables, and what should be the sampling unit. In addition, many serious problems have to be resolved about how to cope with the fact that individual studies are not independent and the analyses are

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

once a suitable sample of experiments has been selected, how to make an overall judgment about what patterns, trends, strengths, and weaknesses characterize the sample. Up until recently, such a review of a body of literature has been an unstructured and highly subjective affair. Understandably, two individuals surveying the same body of literature could, and did, often come up with diametrically opposed conclusions.

As cognitive psychologists have emphasized, the capacity of humans to handle mentally a number of items is severely limited. What constitutes an "item" varies greatly with the structure of the material and the individual's previous familiarity and expertise in a given field of knowledge. Even within his field of speciality, a scientist would have great difficulty in trying to comprehend patterns in over a dozen or so reports without external aids and a systematic procedure.

When the nonparapsychologist critic tries to make sense of a large body of parapsychological literature, he is at a great disadvantage. His critical capacities have not been trained to pick out relevant from irrelevant details in seeking interrelationships. Lacking concrete experience with many of the experimental designs, he is at a decided disadvantage in knowing what things could go wrong and which sorts of controls would be critical. And when the number of separate reports is more than a dozen or so, he cannot be expected to be able to grasp the total picture without help from systematic and quantitative summarization procedures.

Yet, so far as I can tell, only two critical evaluations of "systematic lines" of parapsychological research have ever been carried out with any procedure approximating systematic, explicit, and quantitative guidelines. Both of these were carried out fairly recently. One was by Charles Akers, a former parapsychologist with both experience and publications in the field [74]. The other was by myself, acting as an external critic who accepted the parapsychologists' challenge to fairly evaluate a systematic line of research which they feel represents their strongest case for the repeatable experiment [73], [75].

Akers' Methodological Criticisms of Parapsychology

Akers' methodological evaluation of contemporary parapsychological research represents a landmark in parapsychological criticism. Akers, who holds a Ph.D. degree in Social Psychology, has worked as a parapsychologist in Rhine's laboratory and knows the contemporary scene from the inside.

After a careful selection procedure, Akers arrived at a sample of 54 ESP experiments. These experiments had all been cited in the *Handbook of Parapsychology* or other parapsychological literature as exemplars of the evidential database. The selection was restricted to studies in which significant results had been claimed for a sample of relatively unselected percipients. He excluded unpublished reports, studies which were reported only as abstracts or convention reports, and studies which were exploratory or preliminary to a stronger replication. He also excluded experiments which produced scores in the wrong direction ("psi missing") [74].

> The final sample of 54 experiments is fairly complete. If it is not inclusive, it is at least representative of findings in altered state and personality research.

Akers then screened all his 54 studies sequentially through each of the seven criteria to see how many could pass through all of them. He first looked at how many of the studies used inadequate randomization of the targets. Although he found almost half of the studies used inferior methods to randomize targets he considered this to be a "minor contaminant." In his opinion, such randomization failures as he observed would not be sufficient to account for the above chance results which each of these studies obtained.

Next he looked at the possibility of sensory leakage. For example, in several of the Ganzfeld experiments the agent handled the slide or picture which served as the target. Later the percipient was given that very same target along with some foils and asked to select which item had been the target. In such a situation either inadvertent or deliberate cueing is clearly a possibility. A parapsychologist should not be entitled to claim ESP as the explanation for a successful selection by the percipient under such circumstances. Akers assigned a flaw to any experiment which had this or one of his other categories of possibilities for sensory leakage. As many as 22 of the 54 experiments were cited for having at least one flaw of the sensory leakage kind (some had more than one kind).

In a similar fashion, Akers checked for security problems, recording errors, optional stopping, data selection, inadequate documentation, multiple testing, and some additional flaws of a technical nature. On each criterion, Akers assigned a flaw only if, in his opinion, the defect was sufficient to account for the above chance hitting actually reported [74].

> Results from the 54-experiment survey have demonstrated that there are many alternative explanations for ESP phenomena; the choice is not simply between psi and experimenter fraud.... The numbers of experiments flawed on various grounds were as follows: randomization failures (13), sensory leakage (22), subject cheating (12), recording errors (10), classification or scoring errors (9), statistical errors (12), reporting failures (10).... All told, 85 percent of the experiments were considered flawed (46/54).

In other words, only 8 of the 54 experiments—all of which were selected to be best cases—were free of at least one serious flaw on Akers' criteria. But Akers points out a number of reasons to be concerned about the adequacy of even these "flawless" studies [74].

> In conclusion, there were eight experiments conducted with reasonable care, but none of these could be considered as methodologically strong. When all 54 experiments are considered, it can be stated that the research methods are too weak to establish the existence of a paranormal phenomenon.

Akers' conclusion is especially damaging to the case for psi because he leaned over backwards to give the benefit of doubt to the experimenters. In some cases where the documentation was incomplete, Akers assumed that the investigator had taken the proper precautions against sensory leakage. And Akers did not assign flaws to experiments if their randomization procedures were less than optimal (he considered this to be only a "minor contaminant"). Experiments that were deficient on his other criteria such as optional stopping and others were not assigned flaws if, on Akers' judgment, the deficiency on that criterion was insufficient to have caused the total number of hits. In other words, Akers was not judging whether the experiment had

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

conducted "*post hoc*." Yet, it has many advantages over the previously unstructured and subjective assessments. The reviewer is forced to make many more of his or her standards and procedures explicit. The resulting debate can be more focussed and the specific areas of disagreement can be pinpointed more accurately. In addition, the use of quantitative summaries often brings out patterns and relationships that would ordinarily escape the unaided reviewer's cognitive limits.

Along with an increase in more informed and constructive criticism there are signs that the parapsychological community is responsive and willing to change both its procedures and claims in line with some of the criticisms. Although we still disagree strongly on many of the issues, Honorton has made many changes in his claims and procedures in a sincere effort to take some of my criticisms into account [73], [77]. At its 1984 annual meetings in Dallas, TX, the Parapsychological Association established a committee which will attempt to establish guidelines for the performance of acceptable experiments in various lines of parapsychological research. Along with some major parapsychologists such as Honorton, the committee includes both internal critics such as Akers and external ones such as myself.

My survey of psychical research from the time of Hare and Crookes to the present has suggested that, although the specific evidence put forth to support the existence of psi changes over time, many of the key issues and controversies have remained unchanged. The parapsychologists still employ similar strategems to seemingly enable them to stick to their claims in the face of various incosistencies. And the critics, sharing many assumptions with the proponents, still behave in rather emotional and irrational ways. Indeed, the level of the debate during the preceding 130 years has been an embarrassment for anyone who would like to believe that scholars and scientists adhere to standards of rationality and fair play.

I suspect it is because the quality of the criticism has been so poor and its content so obviously irrelevant that parapsychologists have managed to live so long with the illusion that the quality of their evidence was so much better than it really was. Both Akers and I were surprised to find how defective, in terms of the most elementary standards, the best of the contemporary parapsychological research really was. I know that some parapsychologists have been surprised to realize how far the current status of psi research departs from the professed standards of their field. And I would not be surprised that most of the rest of the parapsychological community, in the absence of systematic and critical surveys, had assumed that their database was of a much higher quality than it, in fact, is.

All this suggests, as I have already indicated, that the parapsychological evidence, despite a history of more than 130 years of inquiry, is not ready to be placed before the scientific community for judgment. The parapsychologists' first order of business should be to get their own house in order. They no longer can safely assume that the typical parapsychologist has the competence to correctly use statistical tools, design appropriate investigations, carry out these investigations correctly, or to write them up properly. Indeed, the evidence suggests the opposite. Both the Parapsychological Association and the parapsychological journals have to establish explicit guidelines and minimal standards. Then they have to make sure that members of their profession become fully aware of these standards and recognize the necessity for living up to them.

References

[1] R. G. Jahn, "The persistent paradox of psychic phenomena: an engineering perspective," *Proc. IEEE*, vol. 70, pp. 136–170, Feb. 1982.

[2] G. Murphy and R. O. Ballou, *William James on Psychical Research*. New York: Viking Press, 1969.

[3] A. Flew, "Parapsychology revisited: laws, miracles and repeatability," in *Philosophy and Parapsychology*, J. Ludwig, Ed. Buffalo, NY: Prometheus Books, pp. 263–269.

[4] R. Hyman, "Pathological science: towards a proper diagnosis and remedy," *Zetetic Scholar*, no. 6, pp. 31–39, July 1980.

[5] R. L. Moore, *In Search of White Crows: Spiritualism, Parapsychology and American Culture*. New York: Oxford Univ. Press, 1977.

[6] F. Podmore, *Mediums of the 19th Century*. New Hyde Park, NY: University Books, 1963 (two volumes).

[7] M. Faraday, "Experimental investigation of table-moving," *The Athenaeum*, pp. 801–803, July 2, 1853.

[8] R. Hare, *Experimental Investigation of the Spirit Manifestations, Demonstrating and Existence of Spirits and their Communion with Mortals: Doctrine of the Spirit World Respecting Heaven, Hell, Morality, and God. Also, the Influence of Scripture on the Morals of Christians*. New York: Partridge & Brittan, 1855.

[9] L. Shepard, Ed., *Encyclopedia of Occultism and Parapsychology*. Detroit, MI: Gale Research Co., 1978.

[10] I. Asimov, *Asimov's Biographical Encyclopedia of Science and Technology*. New York: Equinox, 1976.

[11] W. George, *Biologist Philosopher: A Study of the Life and Writings of Alfred Russel Wallace*. New York: Abelard-Schuman, 1964.

[12] M. J. Kottler, "Alfred Russel Wallace, the origin of man, and spiritualism," *Isis*, vol. 65, pp. 145–192, 1974.

[13] F. M. Turner, *Between Science and Religion: The Reaction to Scientific Naturalism in Late Victorian England*. New Haven, CT: Yale Univ. Press, 1974.

[14] A. R. Wallace, *My Life: A Record of Events and Opinions*. New York: Dodd, Mead, & Co., 1906.

[15] R. G. Medhurst, Ed., *Crookes and the Spirit World: A Collection of Writings by or Concerning the Work of Sir William Crookes, O.M., F.R.S., in the Field of Psychical Research*. New York: Taplinger, 1972.

[16] E. E. F. D'Albe, *The Life of Sir William Crookes*. New York: D. Appleton and Co., 1924.

[17] J. Palfreman, "William Crookes: spiritualism and science," *Ethics Sci. Med.*, vol. 3, pp. 211–227, 1976.

[18] S. J. Davey, "The possibilities of mal-observation and lapse of memory from a practical point of view: Experimental investigation," *Proc. Soc. Psychical Res.*, vol. 4, pp. 405–495, 1887.

[19] E. J. Dingwall, *Very Peculiar People: Portrait Studies in the Queer, the Abnormal and the Uncanny*. New Hyde Park, NY: University Books, 1962.

[20] J. Hasted, *The Metal-Benders*. London, England: Routledge & Kegan Paul Ltd., 1981.

[21] R. Hyman, [Review of *The Geller Papers*], *The Zetetic* [*The Skeptical Inquirer*], vol. 1, pp. 73–80, Fall/Winter 1976.

[22] ——, "Psychics and scientists: 'Mind-Reach' and remote viewing," *The Humanist*, vol. 37, pp. 16–20, May/June 1977.

[23] ——, "Scientists and psychics," in *Science and the Paranormal*, G. O. Abell and B. Singer, Eds. New York: Scribner's, 1981, pp. 119–141.

[24] ——, "Outracing the evidence: the muddled 'Mind Race'." *The Skeptical Inquirer*, vol. 9, pp. 125–145, 1984–1985.

[25] D. Marks and R. Kammann, *The Psychology of the Psychic*. Buffalo, NY: Prometheous Books, 1980.

[26] J. Palfreman, "Between scepticism and credulity: A study of Victorian scientific attitudes to modern spiritualism," in *On*

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the Margins of Science: The Social Construction of Rejected Knowledge*, R. Wallis, Ed. Staffordshire, England: Univ. of Keele, 1979, pp. 201–236.

[27] C. Panati, Ed., *The Geller Papers: Scientific Observations on the Paranormal Powers of Uri Geller.* Boston, MA: Houghton Mifflin, 1976.

[28] R. Targ and H. E. Puthoff, *Mind Reach: Scientists Look at Psychic Ability.* New York: Delacorte, 1977.

[29] J. C. F. Zöllner, *Transcendental Physics.* New York: Arno Press, 1976 [Reprint of 1888 edition published by Colby & Rich, Boston].

[30] H. Sidgwick, [Presidential Address], *Proc. Soc. Psychical Res.*, vol. 1, pp. 7–12, 1882–1883.

[31] J. B. Hasted, D. Bohm, E. W. Bastin, and B. O'Regan, "Experiments on psychokinetic phenomena," in *The Geller Papers*, C. Panati, Ed. Boston, MA: Houghton Mifflin, 1976, pp. 183–196.

[32] A. Gauld, *The Founders of Psychical Research.* New York: Schocken, 1968.

[33] W. F. Barrett, E. Gurney, and F. W. H. Myers, "First report on thought-reading," *Proc. Soc. Psychical Res.*, vol. 1, pp. 13–34, 1882–1883.

[34] E. Gurney, F. W. H. Myers, and W. F. Barrett, "Second report on thought-transference," *Proc. Soc. Psychical Res.*, vol. 1, pp. 70–89, 1882–1883.

[35] E. Gurney, F. W. H. Myers, F. Podmore, and W. F. Barrett, "Third report on thought-transference," *Proc. Soc. Psychical Res.*, vol. 1, pp. 161–181, 1882–1883.

[36] F. Nicol, "The founders of the S. P. R.," *Proc. Soc. Psychical Res.*, vol. 55, pp. 341–367, 1972.

[37] W. F. Barrett, E. Gurney, and F. W. H. Myers, "Thought-reading," *The Nineteenth Century*, vol. 11, pp. 890–900, June 1882.

[38] J. E. Coover, *Experiments in Psychical Research.* Stanford, CA: Stanford Univ. Press, 1917.

[39] E. Gurney, F. W. H. Myers, and F. Podmore, *Phantasms of the Living.* London, England: Trubner's, 1886.

[40] G. S. Hall, [Reviews of the *Proceedings of the English Society for Psychical Research*, July, 1882 to May 1887, and of *Phantasms of the Living*], *Amer. J. Psychol.*, vol. 1, pp. 128–146, 1887–1888.

[41] T. H. Hall, *The Strange Case of Edmund Gurney.* London, England: Duckworth, 1964.

[42] O. Pfungst, *Clever Hans.* New York: Holt, 1911.

[43] R. N. Giere, *Understanding Scientific Reasoning.* New York: Holt, Rinehart & Winston, 1979.

[44] E. Gurney, "Note relating to some published experiments in thought-transference," *Proc. Soc. Psychical Res.*, vol. 5, pp. 269–270, 1898.

[45] C. Tart, [Review of *The Magic of Uri Geller* by the Amazing Randi; *The Geller Papers* edited by Charles Panati; and *My Story* by Uri Geller], *Psychol. Today*, pp. 93–94, July 1976.

[46] J. B. Rhine, *Extra-Sensory Perception.* Boston, MA: Bruce Humphries, 1935 [Originally published in 1934 by the Boston Society for Psychic Research].

[47] S. H. Mauskopf and M. R. McVaugh, *The Elusive Science: Origins of Experimental Psychical Research.* Baltimore, MD: The Johns Hopkins Press, 1980.

[48] B. Wolman, Ed., *Handbook of Parapsychology.* New York: Van Nostrand-Reinhold, 1977.

[49] K. M. Goldney, "The Soal-Goldney experiments with Basil Shackleton (BS): A personal account," *Proc. Soc. Psychical Res.*, vol. 56, pp. 73–84, 1974.

[50] S. G. Soal and F. Bateman, *Modern Experiments in Telepathy*, 2nd. ed. New Haven, CT: Yale Univ. Press, 1954.

[51] K. M. Goldney, "Obituary: Dr. S. G. Soal, M. A., D. Sc.," *J. Soc. Psychical Res.*, vol. 48, pp. 95–98, 1975.

[52] R. A. McConnell, [Review of *Modern Experiments in Telepathy*], *J. Parapsychol.*, vol. 18, pp. 245–258, 1954.

[53] G. R. Price, "Science and the supernatural," *Science*, vol. 122, pp. 359–367, 1955.

[54] S. G. Soal, "On 'Science and the supernatural,'" *Science*, vol. 123, pp. 9–11, 1956.

[55] J. B. Rhine, "Comments on 'Science and the supernatural,'" *Science*, vol. 123, pp. 11–14, 1956.

[56] C. Scott and P. Haskell, "Fresh light on the Shackleton experiments?" *Proc. Soc. Psychical Res.*, vol. 56, pp. 43–72, 1974.

[57] R. G. Medhurst, "The fraudulent experimenter: Professor Hansel's case against psychical research," *J. Soc. Psychical Res.*, vol. 44, pp. 217–232, 1968.

[58] C. E. M. Hansel, *ESP: A Scientific Evaluation.* New York: Scribner's, 1966

[59] ______, *ESP and Parapsychology: A Critical Re-Evaluation.* Buffalo, NY: Prometheus Books, 1980.

[60] B. Markwick, "The Soal-Goldney experiments with Basil Shackleton: New evidence of data manipulation," *Proc. Soc. Psychical Res.*, vol. 56, pp. 250–277, 1978.

[61] W. Broad and N. Wade, *Betrayers of the Truth.* New York: Simon & Schuster, 1982.

[62] J. Hixson, *The Patchwork Mouse.* New York: Doubleday, 1976.

[63] T. Hall, "Florence Cook and William Crookes: A footnote to an enquiry," *Tomorrow*, vol. 11, pp. 341–359, Autumn 1963.

[64] C. Honorton, "Psi and internal attention states: Information retrieval in the ganzfeld," in *Psi and States of Awareness*, B. Shapin and L. Coly, Eds. New York: Parapsychology Foundation, 1978, pp. 79–90.

[65] ______, "Psi and internal attention states," in *Handbook of Parapsychology*, B. B. Wolman, Ed. New York: Van Nostrand-Reinhold, 1977, pp. 435–472.

[66] E. C. May, B. S. Humphrey, and G. S. Humbbard, "Electronic system perturbation techniques," SRI Int., Final Rep., Sept. 30, 1980.

[67] H. E. Puthoff and R. Targ, "A perceptual channel for information transfer over kilometer distances: Historical perspective and recent research," *Proc. IEEE*, vol. 64, pp. 329–354, 1976.

[68] R. Targ and K. Harary, *The Mind Race: Understanding and Using Psychic Abilities.* New York: Villard Books, 1984.

[69] J. G. Pratt, "Prologue to a debate: Some assumptions relevant to research in parapsychology," *J. Amer. Soc. Psychical Res.*, vol. 72, pp. 127–139, 1978.

[70] G. Murphy, "The problem of repeatability in psychical research," *J. Amer. Soc. Psychical Res.*, vol. 65, pp. 3–16, 1971.

[71] C. Honorton, "Beyond the reach of sense: Some comments on C. E. M. Hansel's *ESP and Parapsychology: A Critical Re-Evaluation*," *J. Amer. Soc. Psychical Res.*, vol. 75, pp. 155–166, 1981.

[72] J. Palmer, "Extrasensory perception: Research findings," in *Advances in Parapsychological Research, vol. 2: Extrasensory Perception*, S. Krippner, Ed. New York: Plenum, 1978, pp. 59–243.

[73] R. Hyman, "The Ganzfeld Psi experiment: A critical appraisal," *J. Parapsychol.*, vol. 30, pp. 76–82, 1985.

[74] C. Akers, "Methodological criticisms of parapsychology," in *Advances in Parapsychological Research*, vol. 4, S. Krippner, Ed. Jefferson, NC: McFarland, 1984.

[75] R. Hyman, "Does the Ganzfeld experiment answer the critics' objections?" in *Research in Parapsychology*, W. G. Roll, J. Beloff, and R. A. White, Eds. Metuchen, NJ: Scarecrow Press, 1983, pp. 21–23.

[76] C. Honorton and S. Harper, "Psi-mediated imagery and ideation in an experimental procedure for regulating perceptual input," *J. Amer. Soc. Psychical Res.*, vol. 68, pp. 156–168, 1974.

[77] C. Honorton, "Meta-analysis of psi ganzfeld research: A response to Hyman," *J. Parapsychol.*, vol. 49, pp. 51–91, 1985.

[78] Psychophysical Research Laboratories, *1983 Annual Report*, Princeton, NJ.

[79] R. D. Nelson, B. J. Dunne, and R. G. Jahn, "An REG experiment with large data base capability, III: Operator related anomalies," Tech. Note, School of Eng./Appl. Sci., Princeton Univ., Princeton, NJ, Sept. 1984.

[80] B. J. Dunne, R. G. Jahn, and R. D. Nelson, "Precognitive remote perception," Tech. Note, School of Eng./Appl. Sci., Princeton Univ., Princeton, NJ, Aug. 1983.

[81] G. V. Glass, B. McGaw, and M. L. Smith, *Meta-Analysis in Social Research.* Beverly Hills, CA: Sage Publ., 1981.

[82] L. V. Hedges and I. Olkin, "Analyses, reanalyses, and meta-analysis [Review of *Meta-Analysis in Social Research*]," *Contemp. Educ. Rev.*, vol. 1, pp. 157–165, 1982.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

IV ANOMALOUS-MENTAL-PHENOMENA JOURNAL PUBLICATIONS

The publications in the specialized journals for anomalous mental phenomena date back to the late 1800's and are too numerous to include here. The papers in this section, however, were selected from a large number of excellent papers in the reviewed literature as representative of some of the directions for current research.

The number that appears in the upper right–hand corner of the first page for each publication is keyed to the following descriptions:

16. Ryzl, M., "A Model of Parapsychological Communication," *Journal of Parapsychology,* Vol. 30, pp. 18-30, (1966). If anomalous cognition can increase an individual's chances of correctly guessing the toss of a coin, then by using a redundant coding technique, a communications system can be constructed that is 100% accurate. Ryzl describes one such successful application in this paper.
17. Honorton, C., "Precognition and Real-Time ESP Performance in a Computer Task with an Exceptional Subject," *Journal of Parapsychology,* Vol. 51, pp. 291-319, (December, 1987). Honorton describes a series of experiments with a single subject spanning over 20 years. In this rare circumstance, Honorton finds robust, statistical evidence for precognition that remains stable over that time period.
18. May, E. C., Utts, J. M., Humphrey, B. S., Luke, W. L. W., Frivold, T. J., and Trask, V. V., "Advances in Remote-Viewing Analysis," *Journal of Parapsychology,* Vol. 54, pp. 194-228, (September, 1990). This paper describes an application of fuzzy set theory to the analysis of subjective data. It includes examples of remote viewing and a comparison with "ground" truth.
19. May, E. C., Luke, W. L. W., Trask, V. V., and Frivold, T. J., "Observation of Neuromagnetic Fields in Response to Remote Stimuli," *Proceedings of Presented Papers for the 33rd Annual Convention of the Parapsychological Association,* National 4–H Center, Chevy Chase, MD, pp. 168-135, (August, 1990). This paper is a progress report on the use of magnetoencephalography to investigate the brain's response to anomalous-cognition stimuli. Following techniques commonly used in the research of visual evoked responses of the brain, the authors report a statistically significant effect. The paper includes discussion about possible artifacts.
20. Braud, W., Shafer, D., and Andrews, S., "Electrodermal Correlates of Remote Attention: Autonomic Reactions to an Unseen Gaze," *Proceedings of Presented Papers for the 33rd Annual Convention of the Parapsychological Association,* National 4–H Center, Chevy Chase, MD, pp. 14-28, (August, 1990). Continuing a popular research line at the Mind Science Foundation in San Antonio, Braud *et. al.* describe a careful experiments that appears to confirm statistically that one can "tell," without looking, when someone is staring.
21. Lantz, N. D., May, E. C., and Piantanida, T., "Remote Viewing: From what TIme Frame Does the Information Originate?" *Proceedings of Presented Papers for the 33rd Annual Convention of the Parapsychological Association,* National 4–H Center, Chevy Chase, MD, pp. 138-150, (August, 1990). In this paper, Lantz *et. al.* describe a technically complex experiment to determine the temporal source of anomalous-cognition information. While there was statistically robust evidence for an effect, no temporal dependence was found.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

A MODEL OF PARAPSYCHOLOGICAL COMMUNICATION[1]

By MILAN RYZL

RECENTLY the application of extrasensory perception (ESP) as a means of communication has been discussed more and more. However, the numerous popular articles have not been counterbalanced by an equal number of scientific publications, and it is impossible to decide whether the reason is the sensational exaggeration of relatively scarce material, or whether, for strategic reasons, some of the more recent scientific findings have been suppressed.

Because of the potential importance of the problem it will be interesting to see how the question is being solved in the scientific literature which is accessible. The characteristics of ESP are such that the faculty is predestined to serve as a device for gaining information which is normally inaccessible (6, 17, 18)—whether simply that of objective events in the outer world in general or the interception of coded messages in particular. However, the absolute reliability of ESP has not yet been achieved experimentally; the experimenter cannot as yet rely absolutely on the correctness of the responses secured in experiments.

To appraise the degree of reliability of ESP for obtaining information in practice, it is not its occasional spontaneous manifestations that must be examined, but results of controlled experimentation. The latter do not give such dramatically striking results, it is true; but they are more suitable for the exact evaluation necessary for the application of ESP in conveying information. For this, it would be necessary to be able to tell before the check-up of results, and maybe even without it, whether the ESP response was correct.

Unfortunately, experimental results in ESP tests have not been nearly as reliable as sensory perception. The eye or the camera very quickly picks up a great deal of information, and radio and TV

[1] This paper is an adaptation of an article by Dr. Ryzl which appeared in a Czechoslovakian technical journal, *Sdelovaci Technika* (*Communication Technique*), Vol. 12 (1964), No. 8, pp. 299-302.

signals will transmit a considerable amount in a short time unit. Errors and inaccuracies in these areas have been reduced to such an extent that they can be disregarded in this comparison.

On the other hand, ESP in simple card-calling tests is so much less reliable that it results in only a slight deviation from mean chance expectation. More than that, subjects prove to be very unstable in their ability to demonstrate ESP, their performance being subject to considerable fluctuation—even to the temporary disappearance of ESP, or to its manifestation as psi-missing. If ESP were a technical contrivance for getting information, it would have to be regarded today as very inefficient.

And yet, even a slight deviation above chance does represent the acquisition of information. In order to make practical application of it, it is necessary only:

(1) To secure stable performance in a subject, or to be able to recognize those times when he is actually using his ESP ability.

(2) To so concentrate the amount of information picked up in ESP tests that the knowledge looked for can be deduced with a degree of reliability fixed in advance.

A third condition should be that this be possible without undue cost. In the present study, however, economic aspects will not be taken into account, but only the possible practical application of ESP.

HISTORICAL SURVEY

The first accessible published report of an experiment in which the method employed made it possible to gain an item of information that could be used practically is the report by Foster (4). In this experiment a question was asked which could be answered either by *yes* or *no*. The subjects were given a mixed pack of black (meaning *yes*) and red (meaning *no*) cards in opaque envelopes, and were asked to sort them onto black or red "target areas," the black area indicating *true* and the red area *false*. When the correct answer to a given question (which the subject was to discover by ESP) was *yes*, black cards should have been found on the black area and red cards on the red area. If the correct answer to the question was *nó*, an association of unlike colors would be the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

proper response. Unfortunately, in this experiment (in which the subjects were Indian children), no ESP was shown.

In most laboratory tests so far, card-calling or card-matching tests (with cards enclosed in opaque covers) have been used. In order to employ this method to gain specific information, it is necessary first to work out a code arrangement for translating the desired information into a certain sequence of cards. The ESP result (if ESP occurred) must then be decoded to obtain the desired knowledge.

Such a procedure was the basis for a suggestion made by W. H. Clark (2) to forecast the temperature. Clark suggested that the subject call the order of a pack of ESP cards which would be cut on a future day according to the temperature reading in a given newspaper on that day. By a method Clark suggested, high scoring on target series assigned to a certain temperature could be used as a basis for prediction.

Since ESP, as so far observed in the laboratory, has been too imperfect, it would be necessary for practical usage somehow to concentrate the information carried in every call. To do this, there would have to be a large number of independent calls on every target. These results could then be statistically evaluated, and the unreliability of individual calls thus compensated for.

G. W. Fisk and D. J. West (3) used this procedure in an experiment in which different subjects called the same card and the majority call was considered the call for that target. They failed in their objective, however, for insufficient evidence of ESP was secured.

This "repeated-guessing" technique was also used by R. H. Thouless (16). In addition to checking the "majority vote," as Fisk and West had done, Thouless introduced what he called an "index of preference," by which he proposed to compensate for the fact that subjects prefer certain symbols and have unequal numbers of calls for the various targets. In another article, C. Scott (12) solved some statistical problems raised by Thouless' method.

The repeated-guessing technique was also used by me in experiments with Miss J. K. (8). It proved to yield an increase

in the reliability of ESP calls. However, since the work with this subject was interrupted (for family reasons), it was impossible to develop the method further and to prove that the successful result was repeatable.

Because of the instability of ESP performance, it is necessary to find a way to determine in advance whether ESP is occurring in a given experiment and also to what extent. The subject's introspective statements have not proven to be reliable (7). The first experimenter who was concerned with finding a way to estimate the degree of reliability of ESP responses prior to the check-up was C. E. Stuart (13). Later, R. J. Cadoret (1) adapted a motor form of expressing ESP somewhat similar to that of dowsing. He tried to measure the reliability of ESP responses by having the subject make two parallel sets of calls, using one of them as an "index series" to be checked in advance. It was assumed that the level of scoring on the unchecked series would be similar to that on the checked. The subject was unaware which series would be used as the index. Slightly significant scores were obtained.

Another method for the same purpose was designed by R. Taetzsch (14), who proposed the use of dual-aspect targets; for instance, playing cards, both value and color (5). In such tests, one aspect would be evaluated as an index series, the other as the experimental series. Dual-aspect targets have also been used in a recent experiment by Dr. Schmeidler (11), who also has aimed at finding out the most reliable forms of ESP manifestation. She emphasizes the comparison of various methods of evaluation, either by appraising both aspects together or either aspect separately.

An original contribution is a psi communication system designed (but never put into practice) by Taetzsch (15). It is a device to convey information between two points in space or time with a degree of reliability fixed in advance. As corroborated in a paper including the relevant calculations, it is possible to use even an imperfect ESP faculty to deduce reliable information by the selection of one of two possibilities (white-black, yes-no, etc.). The subject is to make his call by pressing down on one of two buttons. The repeated-guessing technique is used and the result is worked out by a computer on the basis of a program put into it previously.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

This program ensures the required reliability of information obtained. The machine thus determines the number of calls necessary to arrive at a reliable result according to the degree of ESP shown. It does not give an answer until this degree of reliability has been attained. The proposed system also allows for psi-missing by the random insertion of index calls between the regular calls. The index calls are to serve as an empirical check as to whether, in the given series, ESP is being expressed in a positive or a negative manner.

It can be seen, then, that parapsychologists have given considerable attention to the problem of finding a way to make ESP practically applicable. It seems that the difficulty has mainly been the fact that no way has been found to make the ESP performance so perfect and constant that lawful control can be obtained. This is the impression one would get from the studies generally accessible in the scientific press. It has been only in the experiments with the subject P.S., concerning whose ESP faculties the author reported previously (9, 10), that a sufficiently stable ESP response was attained to warrant an attempt at the reliable conveyance of information by ESP with a real hope of success.

Experimental Set-Up for Subject P.S.

The general objective of the experiment, carried out in 1962, was the identification by ESP of five numbers of three digits each. Each of these numbers was to be transmitted as an independent experimental unit. However, the object of the experiment was not actually to attain a practically applicable use of ESP to convey information (this could not yet be done, given the present state of affairs) but to furnish experimental proof that such application of ESP is possible in principle—that information can be conveyed by ESP with the required degree of exactness and reliability. At the time of planning this experimental series, P.S. was giving reliable results in distinguishing two colors (white-green) on cards enclosed in opaque covers, and therefore this technique was adapted to the task of identifying a number by ESP.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030014-4

By means of an arbitrary system, the designated number was coded into a certain sequence of the colors of cards. The covers, with the cards inside, were submitted to the subject for identification repeatedly until an adequate number of calls were accumulated. The data, treated analytically, gave a reliable indication of the color of the uppermost side of the card before actual checking. By means of the code, the number could be identified.

More specifically, the experiment proper was carried out thus: First of all, a code was worked out by which a certain sequence of 10 (white or green) colors was allotted to each one of the three-digit numbers ranging from 000 up to 999. Then an assistant drew a number by lot. By means of the code, he then transformed it into a sequence of 10 colors. He next took 10 opaque covers made of stiff cardboard and marked on the reverse side with capital letters A, B. . . . . K; and into these he put the white-green cards (measuring 105 x 150 mm.) in such a way that the sequence of colors indicated by the target number were uppermost in the cover. The covers, or "envelopes," were then sealed shut.

To these 10 covers the assistant then added another 10 sealed covers marked on the back with small letters a, b. . . . . k, which contained cards placed just the opposite to those marked with capital letters. The purpose of this second set was a double one. First, it represented a parallel, independent test of the same fact as the basic set and could thus be a check on its correctness. Second, this set ensured an equal distribution of target cards. In point of fact, the chosen code for some numbers resulted in an unequally balanced sequence of colors so that if the subject should show a preference for calling one color over the other, a distortion could have resulted.

After taking from the assistant the 20 envelopes representing the selected number, the experimenter added to them another 10 envelopes of the same appearance, marked on the back 1, 2. . . . . 10. These contained randomly selected white-green cards prepared by the experimenter, who kept a record of their order. They were to be checked after the cards had all been called the required number of times, and presumably would indicate the quality of ESP which had operated in the series.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030014-4

All 30 of the envelopes were of the same external appearance; they differed only in the markings on the back, and these were hidden from the subject's sight throughout the experiment. It was decided beforehand that if the cards with the numbers should show insufficient evidence of ESP, the entire series would be cancelled. It proved unnecessary to do this, for the ESP result was sufficiently stable throughout the experiment to give the desired reliability.

After preparing the targets, the experimenter thoroughly shuffled all the envelopes and gave them to the subject to identify the color uppermost in them. The subject did so, the conditions excluding all known possibility of sensory perception. The experimenter kept a record of the subject's call on each individual card as it was presented to him. Then the experimenter again shuffled the envelopes thoroughly and again handed them to the subject to be identified. This procedure was repeated 50 times so that a total of 50 calls was made on each envelope. Thorough shuffling between series ensured that there was a different order each time and that the index targets were randomly mixed with the experimental ones.

These 50 individual calls on each of the 30 covers gave 1500 calls altogether. The 500 index calls were checked first to get an idea of the amount of ESP to be expected on the others. When this proved to be sufficient, the evaluation of the experiment proper was made.

The evaluation of the two sets of experimental calls of 500 each was based on a method which had been worked out earlier in connection with previous experiments with the same subject. The criteria were so chosen that if they were met they would give strong promise that the content of the envelopes could be reliably foretold. The procedure was as follows:

The total of 50 calls on envelopes with capital letters and 50 on corresponding envelopes marked with small letters were evaluated separately from two points of view. First, the ratio of white to green calls on each envelope was computed. This was called "Total Score" (See Fig. 1. The total number of white calls is always given on the left and the total green calls on the right of the column.) Then the consistency of the preference for one color or the other on a certain envelope was checked. The 50 calls on each card were

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Cover No.	Basic Series			1st Revision			2nd Revision			3rd Revision			Final State-ment	Target
	Total Score W-G	Ratio of Outstanding Scores W-G	Conclusion (if reached)	Total Score W-G	Ratio of Outstanding Scores W-G	Conclusion (if reached)	Total Score W-G	Ratio of Outstanding Scores W-G	Conclusion (if reached)	Total Score W-G	Ratio of Outstanding Scores W-G	Conclusion (if reached)		
1	33—17	1—0												W
2	15—35	0—2												G
3	23—27	1—1												W
4	25—25	0—0												G
5	16—34	0—1												G
6	25—25	0—0												W
7	27—23	1—1												G
8	27—23	0—0												W
9	32—18	0—0												W
10	22—28	0—0												G
1a				26—24	1—0									G
2a				31—19	1—0									W
3a				29—21	1—0									W
4a				31—19	1—0									W
5a				19—31	1—1									G
6a				26—24	0—0									G
7a				32—18	1—0									G
8a				27—23	1—0									W
9a				33—17	2—0									W
1b							26—24	1—0						W
2b							15—35	0—2						G
3b							12—38	0—3						G
4b							19—31	0—1						W
1c										14—36	0—2			G
2c										19—31	0—0			G
3c										37—13	3—0			W
A	18—32	0—1		19—31	0—1	g	18—32	0—1	g	14—36	0—2	G	G	G
B	34—16	3—1	w	31—19	1—0	W							W	W
C	35—15	2—0		40—10	3—0	w?	36—14	1—0	W				W	W
D	29—21	1—0		34—16	2—0	w?	31—19	2—0	w	43—7	4—0	W	W	W
E	33—17	2—1		33—17	2—0	g?	30—20	1—0		45—5	5—0	W	W	W
F	27—23	0—0		8—42	0—4	G							G	G
G	29—21	0—0		46—4	4—0	W							W	W
H	15—36	0—2	G										G	G
I	21—29	0—2		8—42	0—4	G							G	G
K	39—11	3—0	w	36—14	3—0	W							W	W
a	27—23	0—0		29—21	2—0		22—28	0—1		27—23	1—0			W
b	17—33	0—1		7—43	0—3	G								G
c	32—18	2—0		31—19	1—0		28—22	0—0						G
d	27—23	0—0		24—26	1—1		16—34	0—2		22—28	0—0			G
e	30—20	1—0		39—11	3—0		21—29	0—0		25—25	0—1			G
f	34—16	1—0		40—10	2—0	W								W
g	26—24	0—1		28—22	0—0									G
h	45—6	5—0	W											W
i	28—22	0—0		37—13	2—0									W
k	23—27	0—0		18—32	0—3									G

Fig. 1. An example of the data from which one of the three-digit numbers (in this case 242) was identified. Under the heading "Basic Series" is the ratio of colors (or "total score") guessed for each envelope, and beside it, the results of the five 10-number groups (called "Ratio of Outstanding Scores" in the text). In the "Conclusion" columns, a capital letter means that the criteria for indication were met; a small letter, that they were nearly met. The revision columns from left to right show the gradual elimination of those cards which have met the specified criteria of certitude until, at the bottom of the Final Statement column, all ten large lettered envelopes are seen to match the correct target arrangement which represents the selected three-digit number.

divided into five consecutive groups of 10 calls each. The ratio of white and green calls was evaluated in each of these five groups separately. But only those groups were counted in which the calls of one color predominated considerably over those of the other, at least to the extent of 8-2, 9-1, or 10-0. These were called "Outstanding Scores." (See "Ratio of Outstanding Scores" in Fig. 1.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

On the left is the number of outstanding scores in which white calls predominated, and on the right, the number in which green calls predominated.) The judgment as to the actual color of the card in the envelope was drawn after the results on the envelopes marked with capitals had been compared with their opposites marked with small letters.

The definite conclusion as to the color of the card in an envelope was not drawn until all of the following conditions had been fulfilled at the same time. These criteria were:

1. The ratio of colors guessed on a given (capital or small lettered) envelope was 35-15 or greater. (Ratios closer to the chance expectation of 25-25 were considered undecided.)

2. The ratio of "outstanding scores" (8-2 and better in each group of 10 calls) pointed clearly in the same direction. Hence the preponderance was on the same side as the majority data, namely, at least 3-0 (or better still 4-1, 4-0, or 5-0). That is, out of five groups of ten calls a ratio of 8-2 or 9-1 or 10-0 appeared at least three times (3-0).

3. The results of Points 1 and 2 were borne out by the agreement of the *opposite* cards (small or capital lettered, depending on Point 1) on which there had to be a majority score no lower than 30-20 and a simultaneous indication in the same direction based on the ratios of "outstanding scores" at a rate of at least 1-0, or better (2-0, 3-1, etc.).

4. In case Point 3 was not fulfilled, as when the result on the opposite card was not sufficiently convincing, it was decided to require a more decisive result in the main set of calls to offset this:

a. On Point 1 there must be a score of at least 40-10 or better.

b. On Point 2 there must be a ratio of "outstanding scores" of the order of either 4-0 or 5-0.

When the first basic series of 50 calls on each envelope had been evaluated, those envelopes which met the above criteria were removed from the pack. The remaining ones which did not meet the criteria were mixed with a corresponding number of numbered index envelopes. They were thoroughly shuffled and given the subject again as in a new experiment.

At each similar repetition series, 50 calls were made on each

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

envelope and the result evaluated as before. Those envelopes which again did not meet the criteria were put through the same process until they did reach the criteria.

When all 10 colors indicating the given number had thus been reliably identified, the number was deduced by the original code. Then the result was compared with the assistant's record of the target number.

Altogether, five independent series were carried out in this manner and five three-digit numbers were thus identified without a single mistake. Figure 1 shows an example of the result from which one of these three-digit numbers was identified. It shows also the way in which the information about the content of the individual envelopes was determined in the successive experimental series until all could be correctly identified.

Discussion

The basic objective of the experiment was to show the possibility of identifying targets by ESP with a precision which could be specified before the check-up. The objective was attained.

A problem remains concerning the efficiency of the method. For the time being, the procedure is uneconomical and cumbersome as compared with other means of communication. Altogether, five three-digit numbers were transmitted; but to do this, it was necessary to make 19,350 single color-calls (of which 11,978 were hits and 7,372 were misses). The average speed on the whole was about 400 calls per hour so that the mere accumulation of the data took some 50 hours (with two persons participating). To this we must also add the time necessary for evaluation of results.

It must be admitted, however, that this great consumption of experimental time was due partly to the fact that the empirically chosen criteria were very strict so as to meet the requirement of extreme reliability in the identifications. But in some cases of application it would be sufficient to use a statistically expressed reliability of identification, which would make the criteria less stringent and reduce the number of necessary calls (given the same level of ESP performance of the subject). Numerous ways of saving us

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

from making superfluous calls and of attaining the required reliability of identification more economically are afforded by applying communication theory.

The application of some method of automatically recognizing the point at which identification could be possible would also considerably reduce the time involved. It can be seen from Figure 1 that the calls on quite a number of envelopes exceeded what was necessary to identify them (e.g., targets H, h in the Basic Series, or targets F, f in the First Revision Series).

To find a way to determine the point when further calls would be unnecessary, as Taetzsch has suggested for his case, would save at least several scores of calls.

Conclusion

This experiment is proof that ESP as a means of communication can be practically applied. The present technique may not be suitable, for it was carried out only as a short-distance experiment. In actual usage, long-distance communication would be called for, especially in situations in which radio communication is impossible. For long distances, of course, further research aimed at securing a sufficient stability of ESP performance under long-distance conditions will be necessary.

This experiment had several useful features: First was the repeated-guessing technique for concentrating information. By this method the necessary data could be accumulated by only one subject rather than many.

In addition to this, the use of the index trials served as an indicator of ESP. Exclusive of them, the two parallel series, which checked and complemented each other, added another advantage. Because of them it was possible to get a good idea in the course of the experiment of the extent to which the majority votes on corresponding covers tended to favor opposite sides.

Even though the technique permitted this preliminary estimation of trends, it did not open any avenues by which the subject could have gotten sensory cues. The index cards and the test cards were so randomized that no reasoning on his part could have been of benefit.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Finally, not the simple majority vote but an outstanding majority was used; and in cases when such a majority was not secured, the calling was repeated until it was. Consistency of calling in addition to an adequate majority was a useful auxiliary criterion.

The criteria were set empirically, and these were reached in the experiment. It well may be that economy of procedure can be improved with the addition of revisions in line with the theory of probability and of information theory.

References

1. Cadoret, R. J. The reliable application of ESP. *J. Parapsychol.*, 1955, 19, 203-27.
2. Clark, W. H. A practical application of precognition. *J. Parapsychol.*, 1958, 22, 216-17.
3. Fisk, G. W., and West, D. J. Towards accurate predictions from ESP data. *J. Soc. psych. Res.*, 1957, 39, 157-62.
4. Foster, A. A. ESP tests with American Indian children. *J. Parapsychol.*, 1943, 7, 94-103.
5. Foster, Esther B. Multiple aspect targets in tests of ESP. *J. Parapsychol.*, 1952, 16, 11-22.
6. Kazhinsky, B. B. *Biologicheskaya radiosvyaz* (*Biological Radio Communication*). Kiev: Publishing House of the Ukrainian Academy of Science, 1962.
7. Rhine, J. B. On the nature and consequences of the unconsciousness of psi. *J. Parapsychol.*, 1958, 22, 175-86.
8. Ryzl, M. Training the psi faculty by hypnosis. *J. Soc. psych. Res.*, 1962, 41, 234-52.
9. Ryzl, M., and Pratt, J. G. A further confirmation of stabilized ESP performance in a selected subject. *J. Parapsychol.*, 1963, 27, 73-83.
10. Ryzl, M., and Ryzlova, J. A case of high-scoring ESP performance in the hypnotic state. *J. Parapsychol.*, 1962, 26, 153-71.
11. Schmeidler, Gertrude R. The accuracy of parapsychological information. *Indian J. Parapsychol.*, 1960, 2, 169-73.
12. Scott, C. An appendix to "The repeated guessing technique." *Int. J. Parapsychol.*, 1960, 2 (No. 3), 37-46.
13. Stuart, C. E. An analysis to determine a test predictive of extra-chance scoring in card-calling tests. *J. Parapsychol.*, 1941, 5, 99-137.
14. Taetzsch, R. Application of statistical quality control techniques to statistical psi control problems. (Abstract.) *J. Parapsychol.*, 1958, 22, 304.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

15. ———. Design of a psi communication system. *Int. J. Parapsychol.*, 1962, 4, 35-70.
16. Thouless, R. H. The repeated guessing technique. *Int. J. Parapsychol.*, 1960, 2 (No. 3), 21-36.
17. Vasiliev, L. L. *Experimentalnye issledovanya myslennevo vnushenya* (*Experimental Research into Mental Suggestion*). Leningrad: Publishing House of the Leningrad University, 1962.
18. ———. *Vnushenye na rasstoyanii* (*Long Distance Suggestion*). Moscow: Gospolitizdat, 1962.

Rybalkova 4
Praha 2, Czechoslovakia

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Journal of Parapsychology, Vol. 51, December 1987

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

PRECOGNITION AND REAL-TIME ESP PERFORMANCE IN A COMPUTER TASK WITH AN EXCEPTIONAL SUBJECT

By Charles Honorton

ABSTRACT: An individual participant, Malcolm Bessent, with a history of success in laboratory precognitive psi tasks, completed 1,000 trials in a computer-based experiment comparing precognition and real-time target modes. A diode-based electronic random number generator (RNG) served as the target source. Target mode was randomly selected at the outset of each 10-trial run and was unknown to Bessent until the completion of each run. Bessent's task was to identify the actual target from a judging pool of four graphic "card" images presented on a computer graphics display.

Based on Bessent's prior research history, two formal hypotheses were tested: (a) Bessent would demonstrate statistically significant hitting in the precognitive target mode, and (b) his precognitive performance would be significantly superior to his performance on real-time targets. Series sample size, methods of analysis, and significance criteria were specified in advance. Both hypotheses were confirmed. Bessent's success rate in the precognitive target mode was 30.4% ($n = 490$, $p = .0039$). This is reliably above the 25% chance level on an estimate of the 95% confidence level, which gives 27.2% as the approximate lower bound. Real-time performance did not exceed chance expectation (25.9% hits; $n = 510$, $p = .34$). The difference between the precognitive and real-time modes was significant ($p = .045$). Exploratory analyses suggest that performance was related to response mode and latency: Significant hitting occurred when Bessent's responses were based on cognitive impressions but not when they were based on feelings or guesses, and he was more accurate on trials in which he took more time to make his response.

Extensive randomness tests document the adequacy of the RNG. The tests include global RNG certification runs testing uniform distribution of RNG byte values ($n = 6 \times 10^6$ bytes) and sequential biases ($n = 8 \times 10^6$ bits), as well as tests of the actual target sequence. An empirical cross-check control, recently advocated by the critic Ray Hyman, in which Bessent's responses for one run were deliberately mismatched against targets intended for another run, also yielded results close to chance expectation.

Various rival hypotheses including sensory cues, faulty randomization, data-handling errors, data-selection bias, multiple analysis, and deception are assessed and found to be inadequate. This is the fourth precognition experiment with Bessent, each involving a different methodology and each yielding a statistically significant outcome. The combined result is highly significant ($z = 5.47$, $p = 2.26 \times 10^{-8}$). It is concluded that the results provide evidence for a communications anomaly involving noninferential precognition.

I am extremely grateful to Malcolm Bessent for once again making his remarkable talents available for this study, and to the late James S. McDonnell, Michael Witunski, and the James S. McDonnell Foundation for their generous financial support of my research over the past 12 years. I also wish to thank George Hansen, Norman Herzberg, and Donald McCarthy for valuable comments on an earlier draft of this paper.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

This report describes an automated, computer-based experiment with an exceptional subject whose prior laboratory testing history indicates superior performance in precognitive as opposed to contemporaneous ("real-time") psi tasks. The present experiment was designed to further assess this participant's precognitive ability and to systematically compare his performance in precognitive and real-time target modes on a blind basis.

Research Participant

The participant in this study was Mr. Malcolm Bessent. Bessent, who resides with his family near London, England, is a businessman who first came to our attention in the late 1960s as a result of a series of ostensibly precognitive dreams. Bessent came to the United States in 1969 to participate in experimental studies of his apparent psi abilities at Maimonides Medical Center.

During the 1969 visit, Bessent was the percipient in two precognitive dream experiments involving the use of EEG-REM monitoring techniques (Krippner, Ullman, & Honorton, 1971; Krippner, Honorton, & Ullman, 1972). Both studies involved eight precognitive dream sessions in which Bessent attempted to dream about salient aspects of target material that would be randomly selected sometime after the nocturnal dream session by Maimonides staff personnel who were blind to Bessent's dream content.

In the first study (Krippner et al., 1971), target experiences consisting of multisensory stimuli were created on the basis of a randomly selected word in Hall and Van de Castle's (1966) book, *The Content Analysis of Dreams,* which contains specific content items and frequencies for 500 male dream reports. The target episode was generated and presented to Bessent upon his final awakening in the morning. Correspondences between dream transcripts for each of the eight nights and the eight target episodes were rated on a blind basis by three independent judges who had no other contact with the study. By chance alone, one would expect that the mean of the judges' ratings would result in one correct target-transcript pairing. The mean of the judges' ratings for the correct target-transcript pairs were higher than those for the incorrect (decoy) ratings for five of the eight nights, a statistically significant result.

The second precognitive dream study (Krippner et al., 1972) also involved eight precognitive dream sessions. The procedure differed from the first study in the following respects: (*a*) the targets were predefined audio-visual episodes involving thematically related slides and sound effects (e.g., slides of birds accompanied by an audiotape of bird calls); (*b*) a complex randomization procedure was used to select an entry point into a random number table to determine the target episode; (*c*) the target that served as a precognitive influence on Bessent's dreams on odd-numbered nights served also as a presleep influence on the subsequent even-numbered nights, allowing comparison of precognitive and presleep sensory influences on Bessent's dream content; and (*d*) the target was selected and presented to Bessent approximately 24 hours after commencement of the precognitive session. The judging procedure was similar to that used in the first study. The mean of the three judges' ratings for the precognitive nights again yielded a significant result with five direct hits. Interestingly, the results for the presleep sensory exposures yielded chance results.

In a third precognition experiment involving a binary random number generator (Honorton, 1971), Bessent attempted to predict which of two colored lamps would light after he pressed the button associated with his choice. The RNG automatically registered the number of hits in each 16-trial run on a digital counter. In 15,360 trials, Bessent obtained a statistically significant success rate of 51.2%.

In addition to the precognition studies, Bessent was the percipient in two studies involving contemporaneous ("real-time") ESP. These included a long-distance real-time dream study (Krippner, Honorton, & Ullman, 1973) and a study of "psychic" readings (Stanford & Palmer, 1973). In the 1973 Krippner et al. study, Bessent, sleeping in the Maimonides Dream Laboratory, attempted to dream about randomly selected target slides that were projected to audiences at concerts by the rock group "The Grateful Dead," held over a six-night period in Port Chester, NY, approximately 45 miles from the Maimonides Laboratory. The results were statistically significant, with four of the six sessions yielding direct hits. In the Stanford and Palmer study, Bessent attempted to provide "psychic" readings for 20 absentee target persons whose hair samples were used as token objects. Blind ratings of the readings intended for each target person were compared with the mean ratings of readings intended for three other target persons. Although an interesting post hoc finding relating to EEG alpha frequency was reported, Bessent's overall success rate was nonsignificant.

Table 1 summarizes Bessent's ESP testing history. Column 1 gives the study reference. Column 2 indicates the apparent psi

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 1
SUMMARY OF PREVIOUS ESP EXPERIMENTS WITH MALCOLM BESSENT

				Prop. hits			
Reference	Mode	Description	Trials	MCE	Obs.	z	p
Krippner et al., 1971	Precog.	REM dream	8	.125	.625	2.70	.0035
Krippner et al., 1972	Precog.	REM dream	8	.125	.625	2.70	.0035
Honorton, 1971	Precog.	Binary RNG	15,360	.500	.512	2.89	.0020
Krippner et al., 1973	Real-time	REM dream	6	.167	.667	2.01	.0222
Stanford & Palmer, 1973	Real-time	"Psychic readings"	20			1.03	.6321

z (overall) = 4.15, $p = 1.68 \times 10^{-5}$, "fail-safe" n = 27 studies
z (precognition) = 4.79, $p = 8 \times 10^{-7}$ "fail-safe" n = 22 studies
z (real-time) = 0.69, p = .244
z_{diff} (precognition vs. real-time) = 2.45, p = .014, two-tailed

Note. Probability estimates for the REM dream studies are exact p's based on the matching distribution. The results for the Stanford/Palmer study were based on a Wilcoxon matched-pairs signed-ranks test comparing the rating assigned to the correct reading to the mean of three control readings. The resulting z score was not reported in the abstract. The author thanks Dr. Rex G. Stanford for providing this information.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

mode (precognition or real-time ESP), and Column 3 gives a brief description of the study. The number of trials and the expected and observed proportion of hits are given in Columns 4–6. Columns 7 and 8 report the z score and one-tailed probability values associated with the observed success rate. An overall estimate of Bessent's success rate was obtained by computing Stouffer z's (Mosteller & Bush, 1954; Rosenthal, 1984) across all of the experiments, and separately for precognition and real-time experiments. The "fail-safe n's" estimate the number of unreported studies with z scores averaging zero that would be required to reduce the reported z scores to nonsignificance (Rosenthal, 1984).

Bessent's overall record of achievement is an impressive one: Statistically significant outcomes were obtained in four of the five experiments, and he was more successful in precognitive than in real-time ESP tasks. All three precognition experiments were significant, and they yielded significantly higher overall z scores than did the experiments with contemporaneously existing targets. It is note-

worthy that the three precognition experiments used very different procedures and methods. The two precognitive dream studies involved different target situations and randomization methods, and the precognitive guessing experiment was radically different from the dream studies. Across the three experiments, the latency between Bessent's response and the generation of the target ranged from a few hundred milliseconds in the RNG study to over 24 hours in the second dream study. That all three experiments yielded similarly significant outcomes provides converging evidence (Lykken, 1968) that Bessent was demonstrating precognition.

Bessent went into business after returning to England. Although he maintained occasional contact with me, he had no further involvement in parapsychological research until the summer of 1987 when a business trip to the United States provided an opportunity for a brief visit to the Psychophysical Research Laboratories (PRL). When asked why he has not been an active participant in psi research in the U.K., Bessent responded, "No one ever asked me."

Bessent was 43 years of age at the time of his visit to PRL in July 1987. Prior to the experiment, he completed a 55-item personal history survey, the Participant Information Form (Psychophysical Research Laboratories, 1983), as well as Form F of the Myers-Briggs Type Indicator (Briggs & Myers, 1957; Myers, 1962) and the Personality Assessment System (Saunders, 1986). The report on Bessent's performance on the Personality Assessment System was not available at the time of this report.

Participant Information Form (PIF). Bessent believes very strongly in the reality of psi and reports personal psi experiences in all four of the traditional categories (telepathy, clairvoyance, precognition, and PK). He has had extensive experience with various mental disciplines including hypnosis, relaxation exercises, EEG and EMG biofeedback training, and meditation. Although he describes his current practice of meditation as sporadic, at various times he has practiced meditation intensively. Bessent reports recalling specific dream content, including lucid dreams, almost every day. He enjoys activities requiring an involvement in fantasy and reports that he frequently loses awareness of his surroundings when he becomes involved in an activity. Bessent rates himself as outgoing and moderately competitive.

Myers-Briggs Type Indicator (MBTI). In Myers-Briggs typology, Bessent is classified as an Extraverted, Intuitive, Thinking, Judging type (ENTJ). The continuous scores for the four scales were: EI = 77, SN = 139, TF = 95, and JP = 99. Because the cutoff for each

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

scale is 100 with standard deviations of approximately 20, the "TJ" classification is equivocal and Bessent could as easily be classified as reflecting the "Feeling/Perceptive" temperament as "Thinking/Judging." Strong extraversion and intuition scores are interpreted as indicating a general orientation toward the external world and a preference toward attending to possibilities over actualities (Myers, 1980). Myers describes intuitives as being "comparatively uninterested in sensory reports of things as they are. Instead, [they] listen for the intuitions that come up from their unconscious with enticing visions of possibilities" (Myers, 1980, p. 57).

In addition to the experiment reported in this paper, Bessent participated in two ongoing group experiments and served as the single subject in an RNG-PK experiment. These experiments will be reported elsewhere. The present experiment was conducted between July 14 and 17, 1987.

METHOD

Hardware and Software System

Computer System

The experiment was run on an Apple][Plus computer equipped with 64 kilobytes of random access memory and two disk drives. The computer is connected to an Okidata Microline u82A printer through a parallel interface and to a 19-inch Toshiba color TV through an RF modulator. The game controller, which provides the participant's sole mode of interaction with the computer, is an Apple game paddle.

Random Number Generator

A *PsiLab][* random number generator (RNG) served as the randomness source for this experiment (Psychophysical Research Laboratories, 1985a). The RNG is a "Bierman-RIPP"-type RNG that has been component tested and modified to PRL specifications. The RNG circuitry is contained on a component board, which plugs into a peripheral expansion slot in the Apple][computer. The board converts the analog noise voltage from two independent avalanche noise diodes into two digitized data bitstreams. The high frequency noise components from each diode are coupled into LM311 comparators, which perform the digitization of the random analog input noise voltages. The outputs of the comparators are 'LS/TTL compatible, and their risetimes are further improved by two successive 'LS/TTL buffers. The bitstreams are latched during periods free of computer-generated EMI and fed into binary dividers to reduce first-order effects (unbalance of 1's or 0's). The two data streams are then combined in an exclusive-or (half-adder) circuit to further reduce first-order effects. The combined random digital data stream is clocked into an eight-bit shift register at a bit rate given by the Apple clock divided by 32. Eight new data bits (a fresh data output byte) are accumulated at a rate of approximately 4 kHz (4,000/sec). Independent bits are available at approximately 32 kHz (32,000/sec). Appropriate filtering, decoupling, and shielding are included on the RNG board. Documentation of the adequacy of the RNG as a source of random numbers is presented below in the Results section.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

ESPerciser

General description. ESPerciser™ is an ESP testing system designed as a hybrid between conventional forced-choice and free-response ESP tasks. It combines the simplicity of evaluation associated with forced-choice card-calling ESP tasks with the diversity of target material, emphasis on the individual trial, and deployment of cognitive strategies such as imagery usually found only in free-response psi tasks. The program enables systematic comparison of target mode (clairvoyance and precognition) and response mode (guessing, feeling, impressions).

The participant's task on each trial is to identify a randomly selected target from a "pack" containing the target and three decoys. There are 24 packs each consisting of four "cards" bearing a wide variety of graphic images including ESP card symbols and other geometric forms, line drawings of people, animals, etc., and image-evoking words. A target pack is selected randomly on each trial without replacement; that is, a different target pack is used for each trial within the run. The order of presentation of the four card-images is randomly determined. At the beginning of each trial, the participant sees the *ESPerciser* "Impression Period Screen" shown in Figure 1.

The participant is free to take as much time as desired to form impressions of the target. When he is ready to view the target pack,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Figure 1. The computer screen during the "impression period."

the participant presses a button on a game controller. The amount of time taken prior to presentation of the target pack (Impression Period Latency) is measured by the computer for each trial. Figure 2 shows the Response Screen and a target pack.

The participant uses a dial on the game controller to point to the card he believes is the target. A button-press then registers his response. The amount of time taken prior to registration of a response (Response Latency) is also measured for each trial. Prior to receiving feedback to the correct target, the participant uses the game controller to indicate whether his response was based on an impression, a feeling, or a guess, defined for the participant as follows:

> *Impression*: Your choice was based on a distinct cognitive impression such as an image or verbal association.
>
> *Feeling*: You had no cognitive impression, but felt drawn to your choice.
>
> *Guess*: Your choice was based neither on an impression nor a feeling. You cannot identify any specific reason for your choice.

Upon registering the type of response, the participant receives feedback to the correct target. Hits are rewarded with the word

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

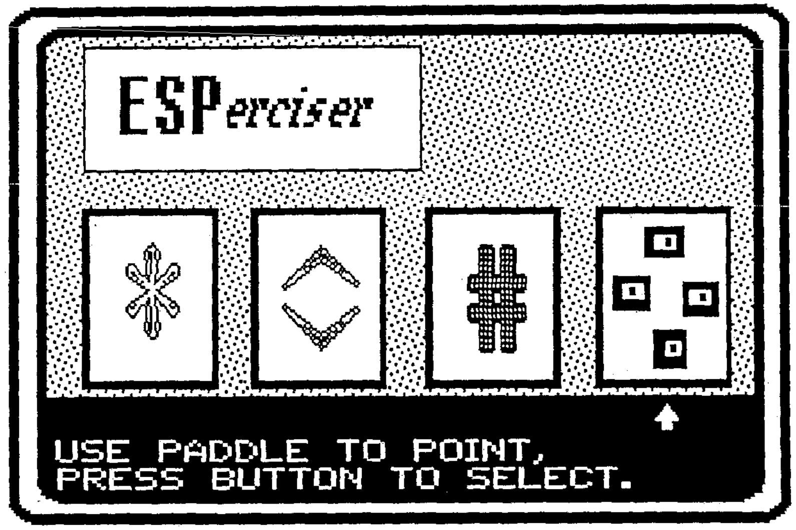


Figure 2. The computer screen during the response period, in which four targets are presented.

"HIT" displayed prominently at the bottom, the screen flashes, and appropriate sound effects are produced.

The above process is repeated for each trial in a run. At the end of each run, the data are saved to disk and the participant is shown a Results Screen indicating the target mode (clairvoyance or precognition), the total number of hits, and a breakdown of trials and hits involving each of the three response types. The participant is then free to initiate another run by using the game controller to respond "Yes" to the prompt "Play again (Y/N)?"

RNG usage. The RNG is used by *ESPerciser* to select the target mode (clairvoyance or precognition) at the beginning of each run. The target pack, pack presentation order, and the actual target are selected through RNG calls on a trial-by-trial basis. Each RNG call is based on a single BASIC PEEK to the RNG slot address. Within the run, *ESPerciser* keeps track of which packs have already been used to insure that a different pack is used on every trial in the run. To reduce positional response biases, the targets are presented in one of four randomly selected orders. In clairvoyance runs, the actual target is selected during the Impression Period, prior to presentation of the four target possibilities. In precognition runs, the target is selected after the participant makes his response and has indicated his basis for the response (guess, feeling, impression).

Security. There is only one legitimate way to exit *ESPerciser* and that is by using the game controller to respond "No" to the prompt "Play again (Y/N)?" (When a participant has completed the maximum number of games specified in the Design module of the Series Manager, described below, the prompt "You have completed the series" is displayed and program control is returned to the Series Manager.) Safeguards built into the program control for the possibility that a participant might accidentally or otherwise illegally terminate execution of the *ESPerciser* program. A flag set at the beginning of each run is written to a special disk file and, upon normal completion of the run, is reset. Any error condition, including an illegal attempt to terminate the program by pressing control-C or control-RESET, or removing the data disk before the data are stored, will cause the computer to "lock up" and trigger a security alarm that can only be deactivated by physically turning off the power to the computer. Any of the above maneuvers leave the error flag set and will cause the Series Manager, when next booted and accessed via the experimenter's private password, to display the message, "THERE IS A DISCREPANCY IN THE SYSTEM! [PARTICIPANT'S NAME] STARTED A GAME OF ESPERCISER AND DID NOT FINISH IT!" This procedure effectively eliminates the following possibilities: (*a*) that an incomplete run could be terminated without detection, (*b*) that program execution could be interrupted to list or alter the program code, (*c*) that someone could prevent data from being properly recorded by either hitting reset at the end of the game or by removing the data disk before data are recorded. If someone replaces the data disk with another to avoid saving "bad" data, that run's data will be missing from the official data disk and their absence will be readily detected. Computer "lock up" also occurs from within the Series Manager if an unauthorized password is entered and fails three times.

File format. The data for each run are saved in a standard ASCII format sequential disk file. The file name is concatenated from the series name, participant ID, run number, and series type. (See following section on Series Manager for description of series type.) Each file includes a header section consisting of the series name, series type, experimenter, participant identification number, participant name, date, run number, target mode, trials per run, total hits, number of trials involving guessing response mode, guessing hits, feeling trials and hits, and impression trials and hits. The individual trial-by-trial data follow the header and include for each trial the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

trial number, participant's response, target, judging sequence, response mode, impression latency, and response latency. After the file is written to disk, it is locked to prevent accidental deletion.

ESPerciser was designed by me and coded in Applesoft® BASIC by Hansen and me. Hansen was also responsible for the development and coding of several 6502 assembler routines, including a binary data compression routine that allows the 96 graphics target card images to be stored compactly on disk.

Series Manager

General description. Series Manager is a password-protected, menu-driven control program that provides the only means through which an experimenter may validly specify parameters for the series design, register new participants in the series, set up a session, and run *ESPerciser*. The Series Manager menu is accessed through entry of a private (and nonechoing) password. As described above, the computer "locks up" and must be rebooted if an unauthorized password is entered and fails three times. The *ESPerciser* Series Manager is an adaptation of Series Manager programs used in the PRL *PsiLab][* and autoganzfeld experiments (Honorton & Schechter, 1987; Psychophysical Research Laboratories, 1985a).

Design. The Design module prompts the experimenter to specify (*a*) the type of series (Pilot, Screening, or Formal); (*b*) the maximum number of runs per participant; (*c*) the maximum number of participants in the series; (*d*) the number of trials per run; and (*e*) the series name. No provision is made for changing the Series Design once accepted by the experimenter, and any subsequent alteration in the design specifications would constitute a deliberate violation of protocol. The design parameters are saved in a disk file and are passed to the *ESPerciser* program at the beginning of each session.

Participant registration. This module prompts the experimenter to input the participant's last name, first initial, and participant identification number. This module verifies that the maximum number of participants specified in Design is not exceeded. (A "SERIES FULL" message appears if an attempt is made to register participants exceeding the Design specification, and control is then returned to the Series Manager menu.)

Directory. The Directory maintains a log of the number of runs completed by each participant and the date of their last session.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Session setup. The Setup module is entered when the experimenter is ready to initiate a session. The participant's name is input by the experimenter, and the module verifies that the participant has been properly registered and has not already completed the maximum number of runs specified by the Series Design.

Run ESPerciser. This module runs an assembly language RNG verification routine to certify the presence of a *PsiLab][* RNG in the Apple computer. (Failure of the RNG verification test results in the message "NO RNG FOUND" and causes the computer to "lock up.") *ESPerciser* is loaded upon verification of the presence of a *PsiLab][* RNG.

The Series Manager was designed by me and was coded in Applesoft BASIC and 6502 assembly language by Rick E. Berger and me. *ESPerciser*-specific adaptations were made by Hansen.

Bessent Design Specifications

The design of this experiment called for Bessent to complete 50 *ESPerciser* runs in each series. The run size was set to 10 trials, and the number of participants per series was set to one. Thus, each series consisted of 500 trials with an expected chance probability on each trial = .25. The series were defined as Formal. Two series (MB1 and MB2) were completed prior to Bessent's return to London. The above design specifications were input by me for each series while Bessent was in another room.

Procedure

Prior to beginning MB1, Bessent was shown *ESPerciser* and was allowed to do several preliminary runs in an informal PRL DEMO series to familiarize himself with the experimental task, use of the game controller to register his responses and response type, the type of feedback provided for hits and misses, and the end-of-run feedback. As with the DEMO series for other PRL computer experiments, the DEMO series was defined as a Screening type series and was used exclusively for the purpose of showing the experiment to prospective new participants and PRL visitors.

Bessent was informed that the target for each trial would be selected by an electronic device in the computer and that his task would sometimes involve "real-time" ESP (i.e., clairvoyance) and at

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

other times would involve precognition. He was given a printed sheet defining the criteria to be used in labeling his responses as Guesses, Feelings, or Impressions. Bessent was informed that mean chance expectation (MCE) was 2.5 hits per run and that his goal was to obtain as many hits as possible on each run. He was encouraged to take as much time as desired on each trial. He was further encouraged to take coffee or cigarette breaks whenever he wished and not to do too many runs at any one time.

Bessent received no technical information concerning the *ESPerciser* or Series Manager programs or the hardware RNG, nor did he at any time ask questions about or otherwise express interest in such matters. He was not informed of the security measures described above, nor was he informed concerning the manner in which data were saved, the type or number of data files, the file formats, and so on. At the time of this experiment, the only written description of *ESPerciser* was a one-page, nontechnical account in the *1985 PRL Annual Report* (Psychophysical Research Laboratories, 1985b), which Bessent had not seen.

Session Procedures

The experiment was conducted in an experimental room, adjacent to the main PRL conference area. There are no windows in the experimental room and only one exit, which opens to the conference area. Bessent sat in a comfortable reclining chair in front of the Apple computer and color TV monitor. At the beginning of each session, the overhead fluorescent lights were turned off and the room was illuminated by a desk lamp adjusted to avoid screen glare.

Bessent got coffee or engaged in conversation with PRL staff in another room while I set up the session. When *ESPerciser* was loaded and ready to run, I escorted Bessent to the experimental room. Bessent and I typically engaged in conversation for a few minutes at the beginning of the session. When Bessent indicated his readiness to begin, I sat silently to his left for the first run or two, occasionally offering words of encouragement. I then left the room while Bessent proceeded with the session. Either I or PRL staff personnel Marta Quant and Linda Moore occasionally came into the room to see how Bessent was doing and to offer encouragement. Such visits were frequent and unannounced. When Bessent was ready to quit for the day, he would come and get me, if I was not

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

already present. I retrieved the *ESPerciser* system disk and the data disk and returned them to my office. Bessent and I then discussed the session and how Bessent thought he had done that day.

Chronology

MB1. The first series was completed in the three-day period between July 14 and 16, 1987. The results were analyzed on the afternoon of July 16.

MB2. The second series was completed in a single two-hour session on Friday, July 17, 1987. Prior to beginning MB2, and while Bessent was out of the experimental room, I changed my password and made two modifications in the *ESPerciser* program. The first modification, made at Bessent's request, caused the program to print a hardcopy of the number of hits at the end of each run. Bessent requested this change so that he could keep track of his performance from run to run. Following initial analysis of MB1, Bessent knew that his precognition performance had been just one hit shy of statistical significance and that his overall success rate was just three hits shy of significance. This was the only request for procedural modification made by Bessent during the course of the experiment. The second program modification changed the file-naming protocol so that shorter filenames (MB01, MB02,... MB50) would be used. This replaced the rather awkward file-naming convention involving concatenation of series name, participant identification number, run number, and series type and was done to expedite transfer of the files from the Apple to an MS-DOS computer at the end of the series. Bessent was not informed of the changes in password and file-naming convention.

Data Handling

Data handling was automated to eliminate any possibility of human recording errors and to facilitate analysis. At the end of each series, the individual data files were transferred from the Apple][to a Leading Edge Model D MS-DOS computer using the Kermit protocol and a serial link between the two computers. A Microsoft QuickBasic program translated the summary header data of the individual run files into records that could be read into a dBASE III Plus database file (BESSENT). A similar program translated the individual trial-by-trial data into a second dBASE file (MBTRIAL), allowing analysis of the target sequence, response mode, success rate by target pack, impression period, and response period delays. Overall hit rate and hit rate by target and response type were analyzed separately for the summary header and trial-by-trial data, and no discrepancies were found to exist between the two files. (It should also be noted that there were no discrepancies in hit count between the disk files and printouts in MB2.)

Statistical Analysis

Sample Size

As stated above, the experimental design called for Bessent to complete 50 runs of 10 trials (i.e., a total of 500 trials) in each series. Two series were completed prior to Bessent's return to London on July 19. There was absolutely no data selection, and the number of series was dependent solely on Bessent's flight schedule, which had been arranged prior to beginning the experiment.

Predictions and Planned Analyses

On the basis of Bessent's prior experimental history, it was predicted that he would demonstrate significant above-chance hitting in the precognitive target mode and that his precognition performance would be significantly superior to his performance in the clairvoyance condition.

The planned analyses were limited to (*a*) exact one-tailed binomial probabilities (with $p = .25$ and $q = .75$) to test overall hitting by target mode, and (*b*) an independent groups *t* test to evaluate the difference between the mean run scores of the two target modes. The alpha level was set to $p = .05$, one-tailed.

Exploratory Analyses

In addition to the formal planned analyses, a number of exploratory analyses were anticipated. I was particularly interested in the relationship between Bessent's success rate and his mode of response. Because of Bessent's earlier success in the free-response dream studies, I expected that he might be particularly successful in responses based on imagery (i.e., cognitive impressions), but since he had also succeeded in a binary guessing task, there was really no adequate basis for formulating a prediction. I also expected that

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Bessent's performance would fluctuate with respect to the varying content and discriminability of the target packs. Once again, there was no specific basis for predictions prior to the experiment.

RESULTS

Randomness Tests

Global RNG Certification Runs

The RNG used in this experiment is PRL *PsiLab][* RNG #36. RNG #36 has undergone extensive global certification runs using the PRL *PsiLab][* Random Analysis Protocol (Psychophysical Research Laboratories, 1985a, pp. 3–43) both prior to and following the present experiment. The certification runs include both frequency and serial analysis of RNG output generated automatically with no one present. The programs used for this purpose were written in Applesoft BASIC and 6502 assembly language by Donald J. McCarthy, Dept. of Mathematics and Computer Science, St. John's University, and were run on the Apple][Plus computer with the *PsiLab* RNG board mounted in peripheral expansion slot #4.

Frequency analysis. FREQUENCY ANALYZER measures the degree to which the byte values produced by the RNG match the expected distribution wherein each value (0–255) is equiprobable. In addition to the full breakdown using all 256 values individually, FREQUENCY ANALYZER examines the effect of grouping adjacent values into 2, 4, and 16 cells. (The present experiment, for example, produced target values between 1 and 4 by dividing the byte into four equiprobable parts.) Chi-square and Kolmogorov-Smirnov (K-S) tests are calculated for each of the four cell breakdowns. The chi-square test examines the extent to which the RNG has produced an equal distribution of values for each cell breakdown. The K-S test (Knuth, 1981) is a "goodness of fit" test and measures the degree to which the observed distribution of chi-square *samples* for each cell breakdown deviates from the theoretical distribution. A total of 300 samples each consisting of 20,000 FREQUENCY ANALYZER trials (i.e., a grand total of 6×10^6 bytes) have been collected on RNG #36. Neither the overall chi-square nor the sample-based K-S tests revealed any consistent departures from the expected distributions.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Serial analysis. SERIALYZER performs a generalized Serial test (Good, 1953, 1957) on the RNG bitstream. This test assesses the extent to which successive bits produced by the RNG are independent, that is, devoid of predictable patterns. This type of analysis is particularly important in experiments such as the present one, in which participants receive trial-by-trial feedback. It is possible that a target generator could appear quite random by the simpler tests for equiprobable frequency of targets yet produce sequential dependencies (e.g., a tendency for 3's to follow 1's) that might lead a participant unconsciously or otherwise to a spuriously significant success rate. SERIALYZER examines the RNG bitstream for serial patterns to a depth of 8 bits. The output of the analysis is a chi-square test for each of the three serial depths examined (2, 4, and 8 bits). As with the frequency analysis, the results are summarized by a Kolmogorov-Smirnov test that assesses their "goodness of fit" to the expected distribution of chi-square values. A Runs Test is also produced for each sample and summarized by a K-S analysis. A total of 400 samples each consisting of 20,000 SERIALYZER trials (i.e., a grand total of 8×10^6 bits) have been collected on RNG #36. Neither the overall chi-square tests nor the sample-based K-S tests revealed any consistent departures from the expected distributions.

Tests of the Experimental Target Sequence

In addition to the global RNG certification runs, the actual target sequence for the Bessent experiment was subjected to a number of randomness tests to insure proper functioning of the RNG during the experiment.

Distribution of targets. A chi-square test of the distribution of targets indicates that the target values (1–4) were uniformly distributed. The χ^2 with 3 *df* is 3.68, $p = .298$.

Tests for sequential dependency. As indicated in the discussion of the global RNG certification runs, the assessment of potential sequential dependencies is particularly important in ESP experiments where participants receive trial-by-trial feedback. The 1,000-trial target sequence was subjected to an autocorrelation analysis using Version 3.0 of the *SYSTAT* statistics package (Wilkinson, 1986). The autocorrelation analysis was performed with lags from 1 to 9 trials. The analysis (Table 2) revealed no evidence of sequential dependencies.

Empirical cross-check control. Hyman has recently advocated empirical cross-checks as a method for assessing target randomness in ESP experiments (Hyman & Honorton, 1986). Such analyses were

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 2
AUTOCORRELATION OF 1,000-TRIAL TARGET SEQUENCE

Lag	Correlation	Standard error
1	.026	.032
2	−.032	.032
3	−.021	.032
4	.044	.032
5	−.041	.032
6	−.026	.032
7	.051	.032
8	.014	.032
9	−.054	.032

frequently performed in conjunction with the early card-guessing experiments (Pratt et al., 1940/1966). The cross-check control was performed by matching Bessent's responses on run n against targets for run $n + 1$. This yielded a nonsignificant hit rate of .260 (z = 0.70, p = .243, one-tailed). The associated effect size (Cohen's h; Cohen, 1977, pp. 179–213) is 0.02. Figure 3 shows the outcome of the cross-check control in the form of a cumulative deviation graph with trend lines indicating the .05, .01, and .0013 alpha levels (i.e., z scores of 1.65, 2.33, and 3.00, respectively).

Summary

The randomness tests and empirical cross-check control indicate that the target generator used in this experiment provides an adequate source of random numbers and was functioning properly during the experiment.

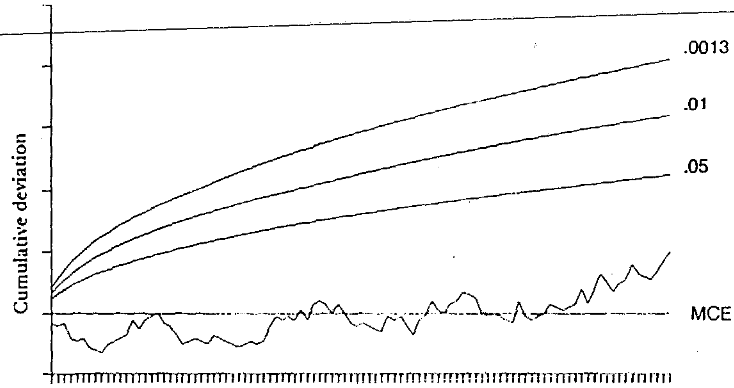


Figure 3. Empirical cross-check control.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

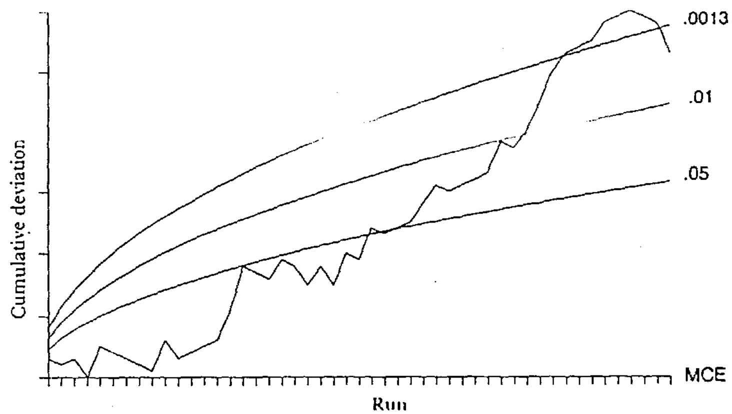


Figure 4. Precognition mode results.

Experimental Results

Planned Analyses

Precognitive target mode. As predicted, Bessent's precognition performance was statistically significant. Forty-nine runs (490 trials) involved the precognitive target mode where the actual target for each trial was randomly selected after registration of Bessent's response. The observed proportion of hits was .304 and is associated with a one-tailed exact binomial p = .0039. This outcome is reliably above the expected chance rate of .25 on a one-tailed 95% confidence interval, which gives .272 as an estimate of the lower limit of the true population mean. The associated effect size is 0.12. Figure 4 shows the outcome of the precognition runs in the form of a cumulative deviation graph.

Clairvoyance target mode. Fifty-one runs (510 trials) involved the clairvoyance target mode in which the target for each trial was randomly selected prior to Bessent's response. They yielded a nonsignificant success rate of .259 (exact binomial p = .34, z = .42). The effect size is .02. Figure 5 shows the outcome of the clairvoyance runs in the form of a cumulative deviation graph.

Table 3 summarizes the results of this experiment by series and target mode.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

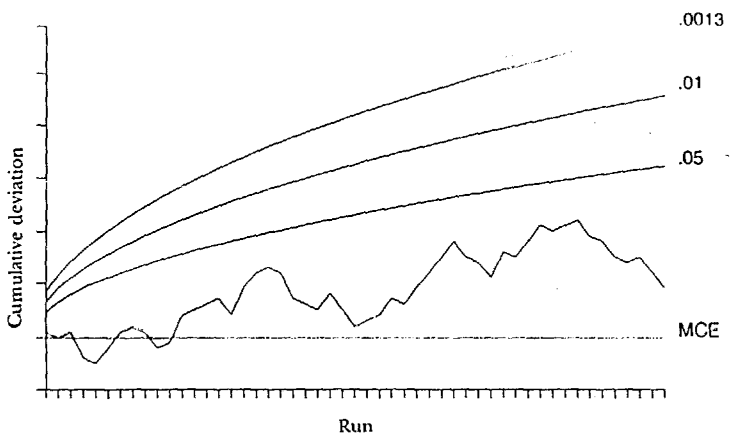


Figure 5. Clairvoyance mode results.

Precognition vs. clairvoyance. Bessent's success rate in the precognition runs was, as predicted, significantly higher than in the clairvoyance runs: $t(98) = 1.71$, $p = .045$, one-tailed. The mean run score for the precognition runs was 3.041 ($SD = 1.353$), and the mean run score for the clairvoyance runs was 2.588 ($SD = 1.299$). The effect size index (r) associated with this difference is .17.

Exploratory Analyses

Overall results. As shown in the last row of Table 3, the overall results of the 1,000-trial experiment, combining the precognition and

TABLE 3
OVERALL RESULTS

Target mode	Trials	Hits	Prop. hits	Run SD	z	p	Effect size (h)
Precognition							
Series 1	270	79	.293	1.49	1.53	.063	.10
Series 2	220	70	.318	1.18	2.21	.013	.15
Total	490	149	.304	1.35	2.66	.0039	.12
Clairvoyance							
Series 1	230	60	.261	1.34	.32	.38	.02
Series 2	280	72	.257	1.29	.23	.41	.02
Total	510	132	.259	1.30	.42	.34	.02
Grand total	1,000	281	.281	1.34	2.20	.014	.07

Note. All p values are exact one-tailed binomial probabilities with $p = .25$ and $q = .75$. z scores are based on the exact p's.

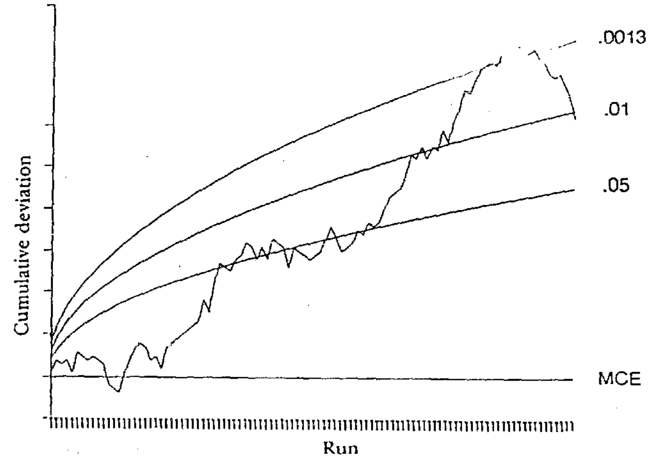


Figure 6. Overall results.

clairvoyance runs, were statistically significant (exact binomial $p = .014$, one-tailed). The overall proportion of hits, .281, is reliably above the expected chance rate of .25 on a one-tailed 95% confidence interval, which gives .258 as an estimate of the lower limit of the true population mean. Figure 6 shows the overall results in the form of a cumulative deviation graph. The distribution of scores is presented in Table 4.

Response mode. As shown in Table 5, Bessent's successful precognitive performance occurred largely on trials in which his responses were labeled as cognitive impressions rather than feelings or guesses. The impression response effect size is approximately double that associated with feeling and guessing responses. Bessent characterized his impression responses as involving fleeting images of shapes. He made responses based on impressions in all of the runs, and they account for approximately two thirds of his responses. Responses based on feelings occurred in 88 of the 100 runs and account for 27% of the total trials. Only 28 runs included responses labeled as guesses, and they account for less than 7% of the trials. Bessent chose "1" a

TABLE 4
RUN SCORE DISTRIBUTION

Target mode	Run score 0	1	2	3	4	5	6	Total
Precognition	1	4	14	13	8	8	1	149
Clairvoyance	2	10	12	13	11	3	0	132
Total	3	14	26	26	19	11	1	281

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 5
RESULTS BY TARGET AND RESPONSE MODE

Target/ response modes	Trials	Hits	Prop. hits	SD	z	p	Effect size (h)
For precognition							
Guess	32	9	.281	.457	.24	.405	.07
Feeling	136	39	.287	.454	.89	.187	.08
Impression	322	101	.314	.465	2.52	.0059	.14
For clairvoyance							
Guess	34	7	.206	.410	−.78	.782	−.11
Feeling	134	37	.276	.449	.61	.271	.06
Impression	342	88	.257	.438	.26	.397	.02

Note. All p values are exact one-tailed binomial probablities with p = .25 and q = .75. The z scores are based on the exact p's.

his response in all 66 of the trials labeled as guesses, indicating that he used the guess category as a null response or "pass" option.

Impression latency. Bessent formed impressions and made his responses quickly. Impression latency, the period from the beginning of the trial to presentation of the target pack, displayed strong variability, and four trials that were extreme outliers were eliminated for the following analyses. Impression latency averaged 5.46 sec/trial (SD = 3.67 sec). Trials involving impression responses were associated with significantly *shorter* impression periods than trials involving feeling or guess responses (Mann-Whitney U Test, z = 2.08, p = .038, two-tailed). Bessent showed a general tendency toward shorter impression periods as the experiment progressed. This is reflected both at the level of the run (Kruskal-Wallis one-way ANOVA by Ranks, $\chi^2[99]$ = 256.4, $p < .001$) and within the run at the trial level (Kruskal-Wallis one-way ANOVA by Ranks, $\chi^2[9]$ = 16.8, p = .052). Impression latency did not vary significantly in relation to hitting or target mode.

Response latency. Bessent's response latency, the period between his initial exposure to the target pack and registration of a response, also showed considerable variability, and the six trials representing the most extreme outliers were removed for the following analyses. Response latency averaged 5.50 sec/trial (SD = .81). In the precognitive target mode, Bessent took significantly *longer* to make his response on trials resulting in hits than on trials resulting in misses (Mann-Whitney U Test, z = 2.58, p = .01, two-tailed). In the clairvoyance target mode, Bessent took slightly more time to make responses re-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

target mode, Bessent took slightly more time to make responses resulting in misses than hits (Mann-Whitney U Test, z = −0.27, p = .787, two-tailed). The correlation between response latency and impression latency was close to zero (r = −.017, $t[992]$ = −.54, p = .589, two-tailed).

Targets. A chi-square test of the distribution of target packs indicates that the packs (1–24) were uniformly distributed: $\chi^2(23)$ = 15.10, p = .891. As described earlier, the 24 target packs contain a wide variety of images ranging from geometric forms such as the standard ESP card symbols to complex drawings of the type commonly used in free-response experiments. Although a one-way ANOVA on the proportion of hits per pack was nonsignificant (F [1,23] = 1.39, p = .251), Bessent's scoring rate ranged from a low of .07 to a high of .43, and the variability of scoring across the 24 packs was suggestively high (Bartlett's Test for Homogeneity of Group Variances, yielding $\chi^2[23]$ = 32.44, p = .091). To assess Bessent's scoring consistency in the two target modes, a Pearson correlation was calculated between the effect size (h) for each pack in the two target modes. The resulting r = .342 ($t[22]$ = 1.71, p = .051, one-tailed) indicates that Bessent exhibited similar scoring patterns on the packs in the two target modes.

Stability of performance. Bessent's precognition mode performance was stable across the two experimental series, and there is no evidence supporting either learning or decline effects. Although inspection of the first two rows of Table 3 indicates that Bessent improved slightly from Series 1 to Series 2, with a 2.5 percent gain in mean hit rate and a suggestive, though nonsignificant, decrease in run score variability (F [26,21] = 1.59, p = .141), the correlation between precognitive run score and run number is close to zero: r = .0496, t (47) = .340, p = .735, two-tailed.

DISCUSSION

Two formal predictions were made prior to beginning the data collection phase of this experiment: (*a*) Bessent would demonstrate statistically significant above chance scoring in the precognition target mode; and (*b*) his precognitive mode performance would be significantly superior to his performance in the clairvoyance target mode. Both predictions were confirmed.

In this section, I consider various rival hypotheses that might account for the experimental outcome, the degree to which the present

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

experiment viewed in conjunction with the earlier Bessent studies constitutes evidence for noninferential precognition, and directions for future research suggested by the exploratory analyses.

Rival Hypotheses

Sensory Cues

Targets in the real-time clairvoyance mode are stored in computer memory, resulting in a small change in the electromagnetic field. Although it is highly unlikely that the human organism is sensitive to such minute fluctuations, this possibility cannot at present be completely excluded. Sensory leakage *is* completely ruled out as an explanation for Bessent's precognitive mode performance, however, because he could hardly gain sensory cues from targets that were selected only after he registered his response.

Faulty Randomization

Five different tests have been presented to document both the general adequacy of the RNG and its proper functioning during the experiment. Extensive global certification runs performed both prior to and following the present experiment indicate that the RNG frequency distribution is well within chance limits and provides no evidence of sequential dependencies. Tests of the target sequence used in the experiment demonstrate that the actual targets were uniformly distributed and revealed no evidence of sequential dependencies. The empirical cross-check control, advocated by Hyman, failed to reveal any tendency for Bessent's responses to coincide with targets for which they were not intended. We conclude that the observed experimental outcome cannot be attributed to faulty randomization.

Data Handling Errors

Automated registration of targets, responses, the checking of hits, and data handling eliminates motivated human checking and recording errors as a basis for explaining the observed experimental outcome. Independent reanalysis of the trial-by-trial target and response data confirmed the summary run hit counts, thus rendering untenable the possibility of machine counting errors.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Data Selection Bias and Multiple Analysis

The sample size parameters (number of participants, runs, and run size) for each series, as well as the Formal status of the series, were specified in advance and stored in a computer disk file as required by the Series Manager protocol. Program error-checking routines insured that these parameters were not exceeded. The number of series, although not specified in advance, was dictated solely on the basis of Bessent's travel plans, which were made prior to the experiment. The second and final series was completed on a Friday morning, and that afternoon was taken up with a visit from the author Arthur M. Young and members of the Princeton University Engineering Anomalies group.

Advance specification of formal predictions and methods of analysis obviates problems involving multiple statistical analyses or multiple indices.

Deception

It is unfortunately true that psi researchers have been occasionally deceived by dishonest participants. Experiments involving a small number of selected individuals are particularly vulnerable to the possibility of subject deception, and when such experiments are successful and are relatively free of more prosaic threats to validity, *suspicion* of subject deception is probably inevitable. For these reasons, I believe this report would be incomplete without some explicit assessment of the likelihood of subject deception.

Even though computer-based experiments such as *ESPerciser* eliminate many known sources of human error, they provide no guarantee against the possibility of deliberate deception. The many media accounts of computer crime and of hackers breaking into supposedly secure government computers attest to this fact. Deception in experiments of this type does, however, require specialized expertise, knowledge of the particular hardware and software systems used, and access to those systems.

In assessing the likelihood of deception, one must consider a number of issues. These include the participant's prior history, motivation, and technical expertise; opportunities for advance or *in situ* preparation; the availability of accomplices; likely methods for deception within a specific research context; and whether there is any *positive evidence* consistent with actual or attempted deception.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

In the 18 years since Bessent first became a research participant in parapsychology, no doubt has ever been raised concerning his integrity. No investigator who has worked with Bessent has ever questioned his reliability or trustworthiness. In working with me, Bessent has always been cooperative and has never attempted to dictate or otherwise exert an influence on experimental protocol; indeed, I cannot recall Bessent ever expressing interest in details of experimental protocol. Bessent is humanistically rather than technically inclined and is rather easily bored by details. According to his own testimony, he has no training or expertise in engineering, computer programming, statistics, or conjuring. I am not aware of anything that would contradict this, and neither I nor other members of the PRL staff observed anything during Bessent's visit that would suggest otherwise.

Bessent visited PRL at his own expense and was not paid for his participation in this experiment. Bessent's best guess, prior to arriving at PRL, was that he would participate in psi ganzfeld experiments and psi experiments "involving computers." He did not receive any detailed information from me prior to his visit regarding current PRL experiments, and, as stated earlier, no detailed description of the *ESPerciser* program existed prior to this report. Besides PRL, Bessent has not visited any other psi research centers that have PRL's *PsiLab][* testing package and might have provided a basis for advance information concerning the RNG or Series Manager program.

During his visit, Bessent was my house guest and had access to PRL only through me. I did not have an Apple computer, *PsiLab][*, or *ESPerciser* software or documentation at home during Bessent's stay. The program and data disks were kept in my PRL office between sessions. Bessent occupied an empty office at PRL and spent his time reading or in conversation with staff when not engaged in the experiment. No experiment-related materials were in the office. At no time was Bessent observed in areas of the laboratory where he had no reason to be.

Opportunities for consorting with potential accomplices were extremely limited. With one exception, Bessent took his meals with me every day during his visit and depended on me for transportation. The exception was July 15, when Bessent went to New York to visit Eileen and Lisette Coly at the Parapsychology Foundation. Bessent had no guests or other visitors during his stay in the Princeton area. Examination of the PRL and my own residence telephone bills during Bessent's visit reveals no out-going calls that are not fully accounted for, and no one at PRL recalls any incoming phone calls for Bessent that were not accounted for. (I have a telephone by my bed and would have been aroused had any calls come at home while I was sleeping.) Bessent did not send or receive any mail that could have contained computer diskettes, program listings, or documentation.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

To account for the observed experimental results on the basis of deception, it would be necessary to assume that either Bessent or an accomplice had extensive knowledge of Applesoft BASIC (and for certain scenarios, 6502 assembly language), access to the *ESPerciser* system disk, and time to learn the code and try different possibilities. That the experiment was conducted over a four-day period constrains the time potentially available for extended tinkering.

The results of the randomness tests conducted on the actual target sequence eliminates any method based on physical intervention with the RNG, for example, forcing the nonrandom occurrence of certain target values or pairings that could later be matched by appropriate responses. If additional data disks with files containing "good" runs had somehow been prepared in advance and later substituted for or copied onto the official data disk, the culprit would have encountered the unexpected change in filenaming format introduced by me without Bessent's knowledge immediately prior to the two-hour session that comprised series MB2. Upon discovery of this change, the culprit would have had to individually unlock, rename, and relock the copied or substituted files. He or she would then have to resort the disk directory so that the files would appear in chronological order, and this would require a special disk utility. To successfully modify data files, a text editor would have been necessary and changes would have had to be made in three separate places for each modified hit count: (*a*) summary hit count, (*b*) summary hits by response type count, and (*c*) the trial-by-trial target or response data. Any substitutions, copying or editing of files, or other attempted manipulations during the experimental session would have had to occur with the knowledge that I or other members of the PRL staff could walk in at any time. Of the 12 experimental runs yielding 5 or more hits, 4 were observed by me and 3 by Quant (including the single run of 6 hits). I was present at the beginning of both series, and Quant was present when Bessent finished series MB2. The *ESPerciser*/Series Manager security system, about which Bessent had not been informed, was not activated at any time, and no procedural irregularities were observed during the course of the experiment.

Precognition

The experimental results provide further support for the existence of a communications anomaly involving noninferential precog-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

nition. This is the fourth experiment involving precognition tasks with Bessent. Each experiment has had a different methodology: Four distinct methods of randomization have been used, including two different, complex procedures for gaining entry points in random number tables and two automated procedures involving different hardware random number generators; the tasks have varied from dream-generated free-response imagery matched against multi-sensory target ensembles by independent blind judges to forced-choice matching of waking state guesses or impressions against machine-selected random numbers represented by lights or images; the latency between response and target selection has varied from a few hundred milliseconds to over 24 hours. All four experiments have yielded statistically significant outcomes, and the combined result is associated with a z score of 5.47 ($p = 2.26 \times 10^{-8}$).

Precognition vs. Real-Time Psi Performance

Bessent was not informed of my prediction that he would display superior performance in the precognition mode, but he was certainly aware of his own prior research history. Nevertheless, Bessent strongly believes his abilities include contemporaneous psi functioning as well as precognition, and his superior performance in precognitive psi tasks is not likely to be based on any conscious bias toward precognition. Such bias could not, in any case, account for his differential performance in the present experiment because the target mode was selected randomly on a run-by-run basis and he was blind to the target mode until the end of each run. It is unclear whether the systematic difference observed between precognitive and clairvoyance tasks reflects a fundamental difference in psi modalities or is linked to individual differences in cognitive style as suggested in Isabelle Myers's (1980, p. 57) discussion of the tendency for MBTI-defined intuitives to be more strongly oriented toward possibilities rather than actualities. Certainly the present study, along with Bessent's prior experimental history, calls into question the generalizability of Tart's (1983) conclusion that precognition does not work as well as real-time ESP and suggests the need for further systematic comparisons of different psi modalities.

Performance and Cognitive Strategies

Bessent was most successful in trials involving cognitive impressions. The effect size achieved on precognitive impression trials was nearly double that obtained on trials based on feelings. (As indicated earlier, Bessent chose to use the "guess" category essentially as a no-response "pass option.") This result is consistent with earlier work involving confidence calls and suggests that participants can to some extent discriminate between internal cues associated with hits and misses (e.g., McCollum & Honorton, 1973). Future research will be necessary to determine whether impressions are generally more accurate than other types of responses or, as seems more likely, that optimal response modalities are related to individual differences in cognitive style.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

References

Briggs, K. C., & Myers, I. B. (1957). *Myers-Briggs Type Indicator Form F.* Palo Alto, CA: Consulting Psychologists Press, Inc.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

Good, I. J. (1953). The serial test for sampling numbers and other tests for randomness. *Proceedings of the Cambridge Philosophical Society*, **49**, 276–284.

Good, I. J. (1957). On the serial test for random sequences. *Annals of Mathematical Statistics*, **28**, 262–264.

Hall, C., & Van De Castle, R. L. (1966). *The content analysis of dreams.* New York: Appleton-Century-Crofts.

Honorton, C. (1971). Automated forced-choice precognition tests with "sensitive." *Journal of the American Society for Psychical Research*, **65**, 476–481.

Honorton, C., & Schechter, E. (1987). Ganzfeld target retrieval with an automated testing system: A model for initial ganzfeld success. In D. H. Weiner & D. I. Radin (Eds.), *Research in parapsychology 1986.* Metuchen, NJ: Scarecrow Press.

Hyman, R., & Honorton, C. (1986). A joint communiqué: The psi ganzfeld controversy. *Journal of Parapsychology*, **50**, 351–364.

Knuth, D. E. (1981). *The art of computer programming: Vol. 2. Seminumerical algorithms* (2nd ed.). Reading, MA: Addison-Wesley.

Krippner, S., Ullman, M., & Honorton, C. (1971). A precognitive dream study with a single subject. *Journal of the American Society for Psychical Research*, **65**, 192–203.

Krippner, S., Honorton, C., & Ullman, M. (1972). A second precognitive dream study with Malcolm Bessent. *Journal of the American Society for Psychical Research*, **66**, 269–279.

Krippner, S., Honorton, C., & Ullman, M. (1973). An experiment in dream telepathy with "The Grateful Dead." *Journal of the American Society of Psychosomatic Dentistry and Medicine*, **20**, 9–17.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Lykken, David T. (1968). Statistical significance in psychological research. *Psychological Bulletin*, **70**, 151–159.

McCollam, E., & Honorton, C. (1973). Effects of feedback on discrimination between correct and incorrect ESP responses: A further replication and extension. *Journal of the American Society for Psychical Research*, **67**, 77–85.

Mosteller, F. M., & Bush, R. R. (1954). Selected quantitative techniques. In G. Lindzey (Ed.), *Handbook of social psychology: Vol. 1. Theory and method* (pp. 289–334). Cambridge, MA: Addison-Wesley.

Myers, I. B. (1962). *The Myers-Briggs type indicator manual.* Palo Alto, CA: Consulting Psychologists Press, Inc.

Myers, I. (1980). *Gifts differing.* Palo Alto, CA: Consulting Psychologists Press.

Pratt, J. G., Rhine, J. B., Stuart, C. E., Smith, B. M., & Greenwood, J. A. (1940/1966). *Extra-sensory perception after sixty years.* Boston: Bruce Humphries.

Psychophysical Research Laboratories (1983). *1983 PRL annual report.* Princeton, NJ: Psychophysical Research Laboratories.

Psychophysical Research Laboratories (1985a). *PsiLab][user's manual.* Princeton, NJ: Psychophysical Research Laboratories.

Psychophysical Research Laboratories (1985b). *1985 PRL annual report.* Princeton, NJ: Psychophysical Research Laboratories.

Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage Publications.

Saunders, D. R. (1986). Predicting psi performance with the Personality Assessment System. In D. H. Weiner & D. I. Radin (Eds.), *Research in parapsychology 1985* (pp. 178–186). Metuchen, NJ: Scarecrow Press.

Stanford, R. G., & Palmer, J. P. (1973). Meditation prior to the ESP task: An EEG study with an outstanding ESP subject. In W. G. Roll, R. L. Morris, & J. D. Morris (Eds.), *Research in parapsychology 1972* (pp. 34–36). Metuchen, NJ: Scarecrow Press.

Tart, C. T. (1983). Information acquisition rates in forced-choice ESP experiments: Precognition does not work as well as present-time ESP. *Journal of the American Society for Psychical Research*, **77**, 293–310.

Wilkinson, L. (1986). *SYSTAT: The system for statistics.* Evanston, IL: SYSTAT, Inc. (Version 3.0).

Psychophysical Research Laboratories
301 College Rd., East
Princeton, NJ 08540

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Journal of Parapsychology, Vol. 54, September 1990

ADVANCES IN REMOTE-VIEWING ANALYSIS

BY EDWIN C. MAY, JESSICA M. UTTS, BEVERLY S. HUMPHREY, WANDA L. W. LUKE, THANE J. FRIVOLD, AND VIRGINIA V. TRASK

ABSTRACT: Fuzzy set technology is applied to the ongoing research question of how to automate the analysis of remote-viewing data. Fuzzy sets were invented to describe, in a formal way, the subjectivity inherent in human reasoning. Applied to remote-viewing analysis, the technique involves a quantitative encoding of target and response material and provides a formal comparison. In this progress report, the accuracy of a response is defined as the percent of the intended target material that is described correctly. The reliability is defined as the percent of the response that was correct. The assessment of the remote-viewing quality is defined as the product of accuracy and reliability, called the figure of merit. The procedure is applied to a test set of six remote-viewing trials. A comparison of the figures of merit with the subjective assessments of 37 independent analysts shows good agreement. The fuzzy set technology is also used to provide a quantitative definition of target orthogonality.

Human analysts are commonly used to evaluate free-response data. Although there are many variations, the basic idea is that an analyst, who is blind to the actual result, is presented with a response and a number of target possibilities, one of which is the intended target. The analyst's task is to decide what is the best response/target match, and frequently includes rank-ordering the targets from best to worst correspondence with the response. It is beyond the scope of this report to provide a critical review of the extensive literature on this topic.

One aspect, however, of this type of evaluation is that analysts are required to make global judgments about the overall match between a complex target (e.g., a photograph of a natural scene) and an equally complex response (e.g., written words and drawings). In a recent book, Dawes (1988) has discussed various decision algorithms in general and the difficulty with global techniques, such as those used in rank-order evaluation, in particular.[1] According to Dawes, the research results suggest that global decisions of this type are not as good as those based on smaller subelements that are later

[1] We are indebted to Professor D. Bem, Cornell University, for directing us to this valuable source of information.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

combined. (See Dawes, 1988, chap. 10, for references to the research.) Humans appear to be capable of deciding what the appropriate variables should be in complex decision processes, but they have proved to be unreliable at combining these variables to arrive at a single decision. Linear algorithms are consistently better at this latter task. Therefore, it seems prudent to develop evaluation techniques that are less sensitive to global decision processes and rely on combinations of more restrictive decisions.

Honorton (1975) has pointed out an additional difficulty inherent in a global rank-order approach. Asking an analyst to rank-order a small set of target possibilities converts the free-response experiment into a forced-choice one, at least on the part of the analyst. It is obvious that in doing so, much quantitative information is lost. For example, a near perfect correspondence between response and target will receive only as much "credit" as one that just barely allowed an analyst to discriminate among the possibilities.

If multiple analysts are used, addition problems arise concerning interanalyst reliability. If an individual analyst judges a number of responses in a series, within-analyst consistency becomes an individual problem.

To address these difficulties, various computer-automated procedures have been suggested in an attempt to reduce the inter-analyst reliability while increasing within-analyst consistency. For examples, see Honorton (1975), Humphrey, May, Trask, and Thomson (1986), Humphrey, May, and Utts (1988), Jahn, Dunne, and Jahn (1980), May (1983), May, Humphrey, and Mathews (1985), and Targ, Puthoff, and May (1977).

In this paper we present the current status of an ongoing research topic. We are not yet ready to propose that the techniques described here be used for free-response analysis; however, we hope to inspire the community to develop a proper set of subvariables so that the problems inherent in global decision processes can be avoided.

Finally, we present a successful application of the mathematical techniques for quantifying target orthogonality for a complex target pool.

Background

Substantial progress has been made in methods for evaluating remote-viewing experiments since the publication of the initial remote-viewing (RV) effort at SRI International (Puthoff & Targ,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

1976). This paper outlines some of the progress and presents the details for one particular method.[2]

Two basic questions are inherent in the analysis of any remote-viewing data, namely, how is the target defined, and how is the response defined.

In a typical outbound RV experiment, definitions of *target* and *response* are particularly difficult to achieve. The protocol for such an experiment dictates that an experimenter travel to some randomly chosen location at a prearranged time; a viewer's task is to describe that location. One method of trying to assess the quality of the RV descriptions in a series of trials is to require that an analyst visit each of the sites and attempt to match responses to them. While standing at a site, the analyst has to determine not only the bounds of the site, but also the site details that are to be included in the analysis. For example, if the target location was the Golden Gate Bridge, the analyst would have to determine whether the buildings of downtown San Francisco, which are clearly and prominently visible from the bridge, were to be considered part of the target. The RV response to the Golden Gate Bridge target could be equally troublesome, because responses of this sort are typically 15 pages of dream-like free associations. A reasonable description of the bridge might be contained in the response; it might be obfuscated, however, by a large amount of unrelated material. How is an analyst to approach this problem of response definition?

The first attempt at SRI at quantitatively defining an RV response involved reducing the raw transcript to a series of declarative statements called concepts (Targ et al., 1977). Initially, it was decided that a coherent concept should not be reduced to its component parts. For example, a *small red VW car* would be considered a single concept rather than four separate concepts, *small, red, VW,* and *car.* Once a transcript had been "conceptualized," the list of concepts constituted, by definition, the RV response. The analyst rated the concept lists against the sites. Although the response was well defined by this method, no attempt was made to define the target site.

In 1982, a procedure was developed to define both the target and response material (May, 1983). It became evident that before a site can be qualified, the overall remote-viewing goal must be clearly defined. If the goal is simply to demonstrate the existence of the

[2] Although the term *remote viewing* is used throughout this paper, the analysis techniques can easily be applied to any free-response data.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

RV phenomenon, then anything that is perceived at the site is important. But if the goal is to gain specific information about the RV process, then possibly specific items at the site are important whereas others remain insignificant.

In 1984, work began on a computerized evaluation procedure (May et al., 1985), which underwent significant expansion and refinement during 1986 (Humphrey et al., 1986). The mathematical formalism underlying this procedure is known as the "figure of merit" (FM) analysis. This method is predicated on descriptor list technology, which represented a significant improvement over earlier "conceptual analysis" techniques, both in terms of "objectifying" the analysis of RV data and in increasing the speed and efficiency with which evaluation can be accomplished. Humphrey's technique, which was based on the pioneering work of Honorton (1975) and its expansion by Jahn, Dunne, and Jahn (1980), was to encode target and response material in accordance with the presence or absence of specific elements.

It became increasingly evident, however, that this particular application of descriptor lists was inadequate in providing discriminators that were "fine" enough to describe a complex target accurately, and unable to exploit fully the more subtle or abstract information content of the RV response. To decrease the granularity of the RV evaluation system, therefore, a new technology would have to allow the analyst a gradation of judgment about target and response features rather than the hard-edged (and rather imprecise) all-or-nothing binary determinations. Requiring an analyst to restrict subjective judgment to single elements rather than to complete responses is consistent with the research reported by Dawes (1988).

A preliminary survey of various disciplines and their evaluation methods (spanning such diverse fields as artificial intelligence, linguistics, and environmental psychology) revealed a branch of mathematics, known as "fuzzy set theory."[3]

Fuzzy Set Concepts

Fuzzy set theory was chosen as the focal point of the RV analytical techniques because it provides a mathematical framework for modeling situations that are inherently imprecise. Because it is such an important component in the analysis, a brief tutorial will be presented to highlight its major concepts.

[3] We wish to thank S. James P. Spottiswoode and D. Graff, CE, for directing us to the fuzzy set literature and for many helpful discussions.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

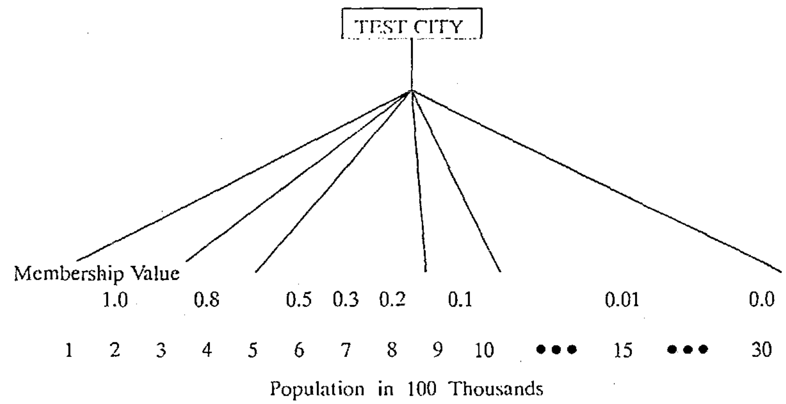


Figure 1. The fuzzy set "kind-of-small" cities.

In traditional set theory (i.e., crisp sets), an element either is or is not a member of a set. For example, the crisp set of cities with population equal to or greater than 1,000,000 includes New York City, but not San Francisco. This set would also *not* include a city with a population of 999,999. The problem is obvious. There is no real difference between cities with populations of 1,000,000 and 999,999, yet one is in the set and the other is not. Humans do not reason this way; therefore, something other than crisp sets is required to capture the subjectivity inherent in RV analysis.

Fuzzy set theory introduces the concept of *degree* of membership. Herein lies the essence of its applicability to the modeling of imprecise concepts. For example, if we consider the size of a city, we might define certain *fuzzy sets*, such as *very small* cities or *kind-of-small* cities. Using *kind-of-small* cities as a fuzzy set example, we might subjectively assert that a city with a population of 100,000 is definitely such a city, but a city with a population of 400,000 is only a little bit like a *kind-of-small* city. As depicted in Figure 1, fuzzy set theory allows us to assign a membership value between 0 and 1 that represents our best subjective estimate as to how much each of the possible city populations embodies the concept *kind-of-small*. In this example, a population of 700,000 assigned a membership value of 0.3.

Clearly, a different set of membership values would be assigned to the populations for the fuzzy sets *very small* cities, *medium* cities, *large* cities, and so forth; a population of 100,000 might receive a value of 0.2 for very small cities, but a value for 1.0 for kind-of-small cities, depending on context, consensus, and the particular

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

application. These membership values can be obtained through consensus opinion, a mathematical formula, or by several other means. Crisp sets are special cases of fuzzy sets, in which all membership values are either zero or one. By using membership values, we are able to provide manipulatable numerical values for imprecise natural language expressions; in addition, we are no longer forced into making inaccurate binary decisions such as, "Is the city of San Francisco large—yes or no?"

In this example, the crisp set of all cities defines the universal set of elements (USE). The crisp set of cities with populations of one million or more is a subset of USE. The fuzzy sets *very small, kind-of-small, medium,* and *large* cities are fuzzy subsets of USE.

Universal Set of Elements

Since targets and the responses will be defined as fuzzy sets, we must specify a USE. The universal set of elements can be quite general and include all aspects of a given target pool, or it can be tailored to a specific experiment to test a given concept (e.g., include only geometric shapes). Since the method of fuzzy set analysis critically depends on the choice of USE, we provide one example that was derived from a target pool used in earlier experiments. What follows is *only* an example of how one might construct a USE. The one we use is not generally applicable to other target pools or other experiments.

We constructed our USE by including a list of features present in photographs from the *National Geographic* magazine with elements obtained from the RV responses in earlier experiments. This USE is presented in Appendix A as the actual coding forms. For the target features, we focused on direct visual elements. (In the general case, other perceptual dimensions can be considered.) In the case of the RV response-derived elements, an effort was made to preserve the vocabulary used by the viewers. Some of the elements, therefore, are either response-dependent or target-dependent or both, whereas others, particularly at the more abstract levels, appear to be more universal across possible USEs.

This universal set of elements is structured in *levels*, ranging from the relatively abstract, information poor (such as vertical lines), to the relatively complex, information rich (such as churches). The current system is structured into seven primary and three secondary levels of elements; the main intent of this structure is to serve as a heuristic device for guiding the analyst into making judicious concrete element assignments based on rather abstract commentary. The use of levels is advantageous in that each element level can be weighted separately and used or not, as the case may be. This enables various combinations of levels to be deployed to identify the optimal mix of concrete versus abstract elements. Of course, any such weighting scheme must be determined in advance of any experiment.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The determination as to which elements belonged on which level was made after consideration of two primary factors: (1) the apparent ability of the viewers to be able to resolve certain features, coupled with (2) the amount of pure information thought to be contained in any given element. Some of these "factor one" determinations were based on the combined anecdotal experience of analysts and monitors in the course of either analyzing or conducting numerous RV experiments; some were determined empirically from post hoc analyses of viewers' abilities to perceive various elements in previous experiments.

The "factor two" determinations were made primarily by arranging the elements such that an element at any given level represents the sum of its constituent elements at lower levels. For example, a *port* element (Level 7) could be considered to include *canal* (Level 6) and *partially bounded expanse of water* (Level 5). The world is not a very crisp place and not all its elements are amenable to hierarchical structuring. Certain violations of the "factor two" rule appear, therefore, throughout the USE example. It should be noted, however, that some of the more glaring violations were largely driven by the "factor one" determinations (i.e., the viewers' abilities to discern certain elements) enumerated above.

To emphasize once again, it is very important to realize that this universal set of elements was constructed to match our particular special targets, viewers, and requirements. They are shown here to illustrate the procedure. Any particular application of fuzzy set technology to the analysis of free-response material requires an a priori construction of an individualized, and improved, USE specific to the target pool and the goals of the experiment.

Target Fuzzy Sets

Each target is defined as a fuzzy set constructed by assigning a membership value to each of the elements in the USE (see Appendix A). In general, membership values can vary continuously on the interval [0,1]. In this application they represent human judgment

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

and, thus, were constrained to vary in steps of 0.1. In addition, they must represent the perceptual dimension used to construct the USE. In our example, membership values were assigned to each element for each of the targets, according to a consensus (on an element-by-element basis) reached by three analysts. This approach was used to mitigate the potential influence of any single coder's biases and idiosyncrasies. A numerical assignment, μ ($0 \leq \mu \leq 1$, in steps of 0.1), was made for each element in response to the following question: How visually important is this element to this photograph?

Encoded by this method, the fuzzy sets served as a formal definition of the targets for the analysis. It should be noted that our USE defined targets in terms of visual importance.[4] If other dimensions are of interest (e.g., conceptual, functional, allegorical), the USE would have to be revised to incorporate them.

In an actual experimental series, it is critical that the target fuzzy sets be defined by analysts *before* the series begins. Because of the potential information leakage owing to bias on the part of the analyst, it is an obvious mistake to attempt to define the target fuzzy set on a target-by-target basis in real time or post hoc.

Response Fuzzy Sets

To define RV response fuzzy sets, membership values μ are assigned for each element in the USE by asking: To what degree am I (the analyst) convinced that this element is represented in this response? For example, if a response explicitly states "water," then the membership value for the water-element should be 1. If, however, the response is a rough sketch of what might be waves, then the membership value for the water-element might be only 0.3, depending on the specificity of the drawing. This definition of membership value is quite general and can be used in most applications.

In our example, responses were coded according to this definition (but still using the USE in Appendix A). The assigned μ's for the targets and responses were one-digit fuzzy numbers on the interval [0,1] (e.g., 0.1, 0.2, 0.3, etc.). In some rare cases, two-digit assignments (e.g., 0.05, 0.15, 0.25, 0.35, etc.) were made; any finer assignments, however, were deemed to be meaningless. Thus, the response was defined as its fuzzy subset of the USE.

[4] Implied visual importance was ignored. For example, in a photograph of the Grand Canyon that did not show the Colorado River, water, river, and so on would be scored as zero. By definition the target was only what was visible in the photograph.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

In an actual experimental series, each response fuzzy set is created by analysts who are blind to the intended target.

Fuzzy Set Definition of Figure of Merit

Once the fuzzy sets that define the target and the response have been specified, the comparison between them to provide a figure of merit (FM) is straightforward. In previous work (Humphrey et al., 1986), we have defined *accuracy* as the percent of the target material that was described correctly by a response. Likewise, we have defined *reliability* (of the viewer) as the percent of the response that was correct. The FM is the product of the two; to obtain a high FM, a response must be a comprehensive description of the target and be devoid of inaccuracies. The mathematical definitions for accuracy and reliability for the *j*th target/response pair are as follows. Let $\mu_k(R_i)$ and $\mu_k(T_j)$ be the membership values for the *k*th element in USE for the *i*th response and the *j*th target, respectively. Then the accuracy and reliability for the *i*th response applied to the *j*th target are given by:

$$\text{accuracy}_{ij} = a_{ij} = \frac{\sum_k W_k \min\{\mu_k(R_i), \mu_k(T_j)\}}{\sum_k W_k \mu_k(T_j)},$$

$$\text{reliability}_{ij} = r_{ij} = \frac{\sum_k W_k \min\{\mu_k(R_i), \mu_k(T_j)\}}{\sum_k W_k \mu_k(R_i)}$$

where the sum over *k* is called the *sigma count* in fuzzy set terminology, and is defined as the sum of the membership values. We have allowed for the possibility of weighting the membership values with weights W_k in order to examine various level/element contributions to the FM. The index, *k*, ranges over the entire USE.

For the above calculation to be meaningful, the μ's for the targets must be similar in meaning to the μ's for the responses. As we noted above, in our definition of the membership values, this is not the case. The target μ's represent the visual importance of the element relative to the scene, and the response μ's represent the degree to which an analyst is convinced that the element is represented in the response regardless of its relevance to that response.

With advanced viewers it might be possible to change the definition of the response μ's to match the definition of the target μ's. In that case, the viewer must not only recognize that an element is

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

present in the target, but must also provide information as to how visually important it is. This ability is currently beyond the skill of most novice viewers. Alternatively, we have opted to modify the target μ definition by using the fuzzy set technique of α-cuts. In our example, an α-cut is a way to set a threshold for visual importance. All target elements possessing that threshold value or higher are considered to be full members of the target set. In fuzzy set parlance, an α-cut converts a fuzzy set to a crisp one. The result is that the target set is now devoid of detailed visual information: a potential target element is either present or absent in the target set, regardless of its actual visual importance. Even with this conceptual change in the target definition, the FM formalism described above remains applicable, because a crisp set can be considered as a fuzzy set with all membership values equal to 0 or 1. It is important to recognize that the α-cut is only applied to the target set; the response set remains fuzzy.

Assessment of Quality of the Remote Viewing

It is difficult to arrive at a general assessment of how well a given response matches a specified target. The ideal situation is to obtain some absolute measure of goodness of match. Although the FM is an approximation to this measure, it is impossible to assess the likelihood of a particular FM value because it requires knowledge of the viewer's *specific* response bias for the session. It is possible to determine general response biases (May et al., 1985), but that knowledge is only useful on the average. For example, a viewer may love rock climbing and may spend most of his free time involved in that activity. Thus, the general response bias would probably entail aspects of mountains, rocks, ropes, and so forth. Suppose, however, that the viewer spent the evening previous to a given RV session on a romantic moonlight sail on San Francisco Bay. For this specific RV session, the response bias might include romantic images of the moonlit water, lights of the city, and bridges.

The current solution to the problem is to provide a *relative* assessment of FM likelihood. A relative assessment addresses the following question: "How good is the response matched against its intended target, when compared to all possible targets that could have been chosen for the session?" This is not ideal, since the answer depends on the nature of the remaining targets in the pool. An example of the worst-case scenario illustrates the problem. Suppose

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

that the target pool consisted of 100 photographs of waterfalls, and the viewer gave a near-perfect description of a waterfall. (We assume that this description is not fortuitous.) An absolute assessment of the resulting FM should be good, whereas a relative assessment will be low. The worst-case scenario can be avoided, to a large degree, by carefully selecting the target pool. (See the later section "A Quantitative Definition of Target Orthogonality.")

To provide a relative assessment of the likelihood of a given FM, we define the score for one session to be the number of targets, n, out of a total, N, that have an FM equal to or higher than the FM achieved by the correct match.[5] The answer to the question: "Given this response, what is the probability of selecting a target that would match it as well as or better than the target selected?" is n/N.

Consecutive RV responses by the same viewer are not statistically independent, nor can the responses be considered to be random in any sense. The statistically independent random element in the session is the target. Since targets are selected with replacement, under the null hypothesis of no psi, the collection of scores derived over a series of m trials constitutes a set of independent random variables, each with a discrete uniform distribution. Under the null hypothesis, the mean chance expectation for the score in each session is given by $(N + 1)/2$ and the variance is given by $(N^2 - 1)/12$. If K is the sum of scores from a series of remote viewings, then the probability of K, under the null hypothesis, can be obtained from the exact distribution for the sum of ranks given by Solfvin, Kelly, and Burdick (1978):

$$p(K \text{ or less}) = \frac{1}{N^m}\sum_{a=m}^{K}\sum_{b=0}^{N}(-1)^b\binom{m}{b}\binom{a-bN-1}{m-1}. \qquad (1)$$

If m is large, then the sum-of-ranks distribution is approximately normal and K/m has a mean of $(N + 1)/2$ and a variance of $(N^2 - 1)/12m$. Thus, a z score can be computed from:

$$z(K \text{ or less}) = \frac{0.5(N+1) - \dfrac{K}{m}}{\sqrt{\dfrac{N^2-1}{12m}}}. \qquad (2)$$

[5] N must be the size of the target pool from which each target was randomly selected, and for this theoretical discussion, we assume no ties.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Ground Truth

To determine whether the new analytical approach was effective, a standard had to be developed against which it could be measured. It was determined that this standard—known as "ground truth"—should consist of a "real-world" normalized consensus about the degree of correspondence between RV responses and their intended targets.

To achieve this objective, we presented analysts (chosen from the general SRI staff) with the same test case of six remote-viewing responses and their associated targets. The test case was the data from a single viewer (177) taken from an experimental series in a 1986 photomultiplier tube experiment (Hubbard, May, & Frivold, 1987). The responses (i.e., two to five pages of rudimentary drawings with some associated descriptive words) were fairly typical of novice viewer output and represented a broad range of response quality. The targets consisted of six photographs of outdoor scenes selected from a *National Geographic* magazine target pool of 200. Thus, this data set was ideally suited for an analysis testbed. Appendix B contains the "best" and "worst" trials (Sessions 9005 and 9004, respectively) from this series in the form of their responses, their intended targets, and their fuzzy set encodings (see the next section).

Each analyst was asked individually for his subjective judgment about the degree of correspondence between the remote-viewing responses and their respective intended targets. The "degree of correspondence" was purposely undefined; the analysts had to formulate their own criteria. The only information provided was that responses typically begin with small bits of information and eventually culminate in a composite drawing at the end. Appendix C contains the coding form that was used to obtain "ground truth."

Each analyst was instructed to examine all of the responses and their intended targets. Then, on a session-by-session basis, he was asked: (1) to assess the degree of correspondence between the remote-viewing response and its intended target, and (2) to register this correspondence assessment by making a vertical hash mark across a 10-cm scale ranging from "none" to "complete."

To perform the ground truth analysis, distance measurements were taken from the left end point of each scale to the vertical slash mark for each assessment. Let the distance obtained for the *k*th session from the *j*th analyst be given by $x_{j,k}$. To account for analysts' biases, the $x_{j,k}$ were normalized by a *z* transformation,

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

$$z_{j,k} = \frac{x_{j,k} - \mu_j}{\sigma_j},$$

where μ_j and σ_j are the mean and standard deviation of the *j*th analyst's distance scores, $x_{j,k}$. The effect of this transformation is to convert an analyst's absolute subjective opinion to a relative one. For the *j*th analyst, the largest $z_{j,k}$ indicates that the degree of correspondence for response/target *k* is higher than *any other pair* in the series. It does *not* indicate overall quality. This type of transformation was necessary since we wished to combine the assessments from a number of different analysts.

To combine the assessments across analysts, we computed the mean *z* score for each response/target pair, *k*, as:

$$z_k = \frac{1}{N_a} \sum_{j=1}^{N_a} z_{j,k},$$

where N_a is the number of analysts. The number of analysts was determined by the data. For the best response/target pair (i.e., session 9005, $k = 5$) we computed the percent change of z_5 for every additional analyst. When the addition of two new analysts produced consecutive changes of less than 2%, the process was considered complete. For this data set, 37 analysts were required before this condition was met. Figure 2 shows the normalized mean for each target/response pair, and represents a relative assessment of remote-viewing quality. These means constitute the basis for the ground truth against which the fuzzy set technique was measured. We recognize that this definition of ground truth is based on global decisions and may not be most optimal (Dawes, 1988).

Results of the Fuzzy Set Analysis

To effect a meaningful comparison between ground truth and the figure of merit analysis, we also analyzed the same RV series that served as the ground truth set by the fuzzy set figure of merit method. The fuzzy set membership values (μ's) for the six targets and six responses were consensus coded by five analysts ranging from expert to novice. A typical spread of μ assignments was $\pm$ 0.1 with an occasional outlier. Some of the elements were vigorously debated until a consensus was reached. Accuracies, reliabilities, and

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

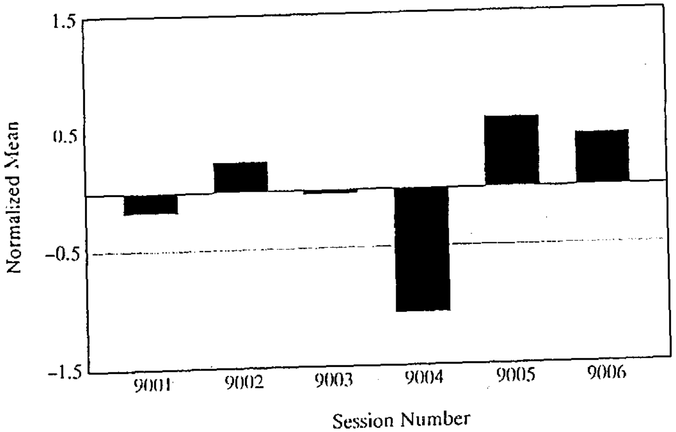


Figure 2. Normalized mean for each target/response pair.

figures of merit were calculated for each target/response pair (Table 1). It should be noted that the encoding was a post hoc exercise, but because the assignment for each element in the USE had to be defended before a consensus was reached, the FMs shown in Table 1 constitute reasonable estimates of their "blind" equivalents. Appendix B shows the target and response elements that were scored from the universal set (see Appendix A) for Sessions 9004 and 9005. As an example of the fuzzy calculation, Appendix B also shows the re-

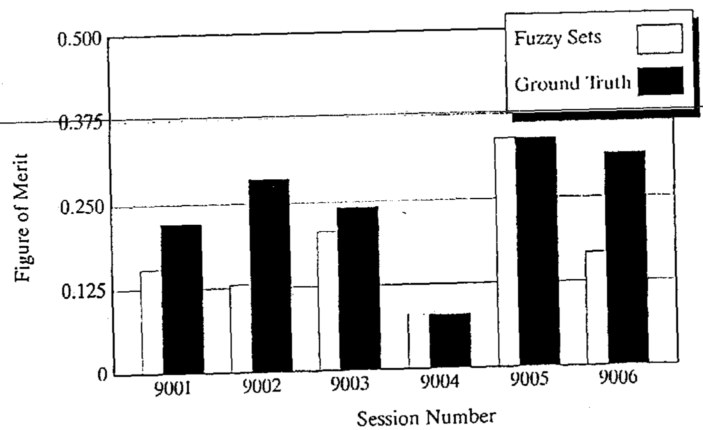


Figure 3. Comparison with ground truth.

TABLE 1
FUZZY SET QUANTITIES FOR "GROUND TRUTH" SERIES

Session	Accuracy	Reliability	Figure of merit	Rank	Fractional rank
9001	.317	.484	.153	80	.403
9002	.273	.477	.130	103	.515
9003	.358	.571	.205	31	.155
9004	.212	.379	.080	142	.713
9005	.573	.594	.340	3	.015
9006	.298	.555	.165	13	.068

sults of the target α-cut, the fuzzy intersection, and the accuracy, reliability, and figure of merit for Session 9005. Table 1 also shows the absolute and relative ranks from a target pool of 200. To determine the absolute rank for each session, we calculated figures of merit for all 200 targets in the pool and placed them in numerical order from the largest to the smallest. The absolute rank is just the position (from the top) of the FM corresponding to the intended target. Ties were resolved by choosing the next larger integer rank number to the centroid of the ties. The fractional rank number can be considered a p value for an individual session and is equal to the absolute rank/200. Using Equation 1, the overall p value for the combined six trials is .052 (N = 200, K = 372, m = 6). Using the approximation (Equation 2), we compute z = 1.633, $p \leq .05$, to demonstrate that for six trials, the approximation in reasonable. For completeness, we compute the effect size (r = 0.67).

To compare the results of the fuzzy set analysis with those of the ground truth, we linearly renormalized the ground truth figures to be within the interval [0,1] and to possess the same maximum and minimum. As can be seen from Figure 3, the results from the fuzzy set analysis system parallel those obtained by a consensus of the 37 analysts each making a subjective assessment of the matches.

These results imply that the combination of (1) the structure of the USE (i.e., the linguistic hierarchical structure), (2) the fuzzy set mathematics, and (3) a consensus approach to assessing the fuzzy sets themselves provided a reasonable representation of the subjective scoring of the same data by a large number of individuals.

A Quantitative Definition of Target Orthogonality

It is often of interest to define how similar or dissimilar targets are to each other. For example, free-response experiments like the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4
Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

ganzfeld often use target packets, with the unselected targets in a packet serving as decoys for judging. Assigning potential targets to packets would be easier with some measure of target orthogonality.

Target definition for the purposes of this mode of analysis is exactly the same as the one described (i.e., a given target is defined by its fuzzy subset of the USE, which has been coded to reflect the visual importance of each target element). The average number of elements, of the total of 131, that was assigned a nonzero value for the targets in our pool of 200 was approximately 37, indicating that the fuzzy set representation of the target pool is rich in visual information. We used this information to determine the degree to which the target set contains visually similar targets.

It is beyond the scope of this paper to describe the extensive work in the literature seeking to find algorithmic techniques that mimic human assessments of visual similarity. One recent article describes techniques similar to the one we used (Zick, Carlstein, & Budescu, 1987).

We begin by defining the simlarity between target i and target j to be a normalized fuzzy set intersection between the two target sets:

$$S_{ij} = \frac{\left(\sum_k W_k \min\{\mu_k(T_i), \mu_k(T_j)\}\right)^2}{\sum_k W_k \mu_k(T_i) \sum_k W_k \mu_k(T_j)},$$

where the index k ranges over the entire USE. We have allowed for the possibility of weighting the membership values with weights W_k to examine various level/element contributions to the target similarities.

For N targets, there are $N(N-1)/2$ unique values (19,900 for $N = 200$) of S_{ij}. The values i and j that correspond to the largest value of S_{ij} represent the two targets that "look" most similar. Suppose another target m is chosen and $S_{m,i}$ and S_{mj} are computed. If both of these values are larger than $S_{m,n}$ (for all n not equal to i or j), then target m is assessed to be most similar to the pair ij. The process of grouping targets based on these similarities is called *cluster analysis*.

Using this process, 200 targets were grouped into 19 clusters, such that the targets are similar within a cluster, and dissimilar between clusters. Table 2 provides an overview of the 19 clusters found from the total analysis of the 200 targets. Some of the names appear to be quite similar, but, in fact, these sets are visually quite distinctive. Figure 4 shows the graphic output of a single cluster in

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 2
NAMES OF THE 19 CLUSTERS

No.	Name	No.	Name
1	Flat towns	11	Cities with prominent geometries
2	Waterfalls	12	Snowy mountains
3	Mountain towns	13	Valleys with rivers
4	Cities with prominent structure	14	Meandering rivers
5	Cities on water	15	Alpine scenes
6	Desert/water interfaces	16	Outposts in snowy mountains
7	Deserts	17	Islands
8	Dry ruins	18	Verdant ruins
9	Towns on water	19	Agricultural scenes
10	Outposts on water		

detail. A much more complex—and visually difficult to understand—graph is generated for the full cluster analysis and is not included here; this smaller subset, therefore, has been chosen to be illustrative of the whole analysis. All targets in this particular sample cluster are islands; the island in each photograph is visible in its entirety. Except for one outlier (i.e., a hexagonal building covering an island), the islands fall into two main groups (i.e., with and without

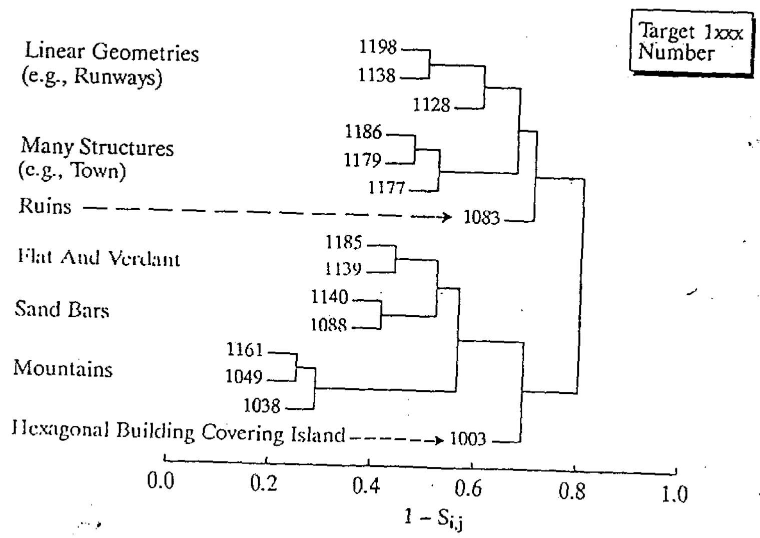


Figure 4. Cluster analysis of island targets.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

manmade elements). The natural islands include three similar mountain islands, two sandbars, and two flat verdant islands.

Using cluster analysis in conjunction with fuzzy set analysis provides for a quantitative definition of sets of targets that are similar to each other within a cluster, but visually different across clusters. Orthogonal clusters can be used to provide visual decoy targets for traditional rank-order judging.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Recommendations and Conclusions

To apply the analysis in its present form to a long RV series is quite labor intensive and, from the results shown in Figure 3, is most likely not justified since this fuzzy set technique approximates human assessment. As we stated in the introduction, however, we are providing only a progress report of ongoing research. Because of the decision concepts described in Dawes (1988) and the obvious benefits of an automated evaluation system, the effort to improve what was described in this paper is certainly justified. The procedure can be used "as is" to improve and quantify target orthogonality.

Several future research areas are suggested to improve the techniques described in this paper. The use of both inter- and intra-level weighting factors needs to be examined systematically. In the analysis described above, all levels and elements were accorded equal weight. The ideal goal would be to determine the optimal weighted mix of abstract versus concrete elements, as a means to achieving the following objectives:

1. Refinement of the cluster analysis for targets, in an effort to simulate, as closely as possible, what is meant by "visual similarities" between targets.

2. Refinement of the analysis of responses, in an effort to achieve even greater correlations between the fuzzy set figure of merit analysis and various forms of ground truth.

Another area that requires examination in some detail is the USE and the hierarchical nature of its structure. It is probable that some elements are more appropriate than others; furthermore, they might be more effectively structured in a semantic network as opposed to a true hierarchy. If a hierarchical structure is retained, then some attention must be paid to the formulation of logical consistency rules that govern element use. This would include numerical relationships governing the membership values (μ's) of higher-order elements (e.g., *port*) vis-à-vis the combined value of their constituent parts (e.g., *city, river, boats, jetties, commercial*).

One inadequacy of the system is that it atomizes conceptual "units." For example, if the response element is *red box*, information is lost in separating *red* from *box*. Current research in fuzzy set theory indicates that fuzzy aggregates of fuzzy elements—"fuzzy sets of fuzzy sets"—are mathematically complex but possible. Some effort should be made to determine whether this technology could be implemented as a means to capturing the information content of the RV response with greater accuracy.

For the visual analysis, research into visual similarities between pictures of natural scenes may serve as a potential refinement tool. The aim here would be to enhance the visual orthogonality of rank-order analysis decoy targets as much as possible. Experiments in normal perception of similarities would assist in determining whether scenes are perceived as similar because of their low-level geometries, concrete elements, or some combination of factors. The ultimate aim would be to refine the target cluster analysis such that it closely simulates ground truth representations of orthogonality.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Appendix A
Coding Forms for the Universal Set of Elements

The following coding forms illustrate the use of a universal set of elements (USE) that matched our particular special targets, viewers, and requirements. We constructed our USE by including a list of features present in photographs from the *National Geographic* with elements obtained from the remote-viewing responses in earlier experiments.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

CONCRETE DESCRIPTOR LEVELS I

Experiment: ______
Trial: ______
Resp./Targ: ______
Coder: ______
Viewer: ______

LEVEL	SINGLE STRUCTURES	SUBSTRUCTURES
10	1 ☐ fort 2 ☐ castle 3 ☐ palace 4 ☐ church (other religious buildings, monastery) 5 ☐ mosque 6 ☐ pagoda 7 ☐ coliseum (stadium, amphitheater, arena)	
9	8 ☐ bridge 9 ☐ [dam (lock, spillway)]	10 ☐ boats (barges) 11 ☐ pier (jetty) 12 ☐ [motorized vehicles (cars, trucks, trains)] 13 ☐ column 14 ☐ spire (minaret, tower) 15 ☐ fountain 16 ☐ fence 17 ☐ arch 18 ☐ wall (e.g., the Great Wall) 19 ☐ monument
8		20 ☐ roads

Figure A1. Coding form.

CONCRETE DESCRIPTOR LEVELS II

Experiment: ______
Trial: ______
Resp./Targ: ______
Coder: ______
Viewer: ______

LEVEL	SETTLEMENT	ELEVATION	LAND/WATER INTERFACE	NO WATER OR VEGETATION	VEGETATION	AMBIENCE/ FUNCTION
7			21 ☐ port (harbor) 22 ☐ [oasis]		23 ☐ agricultural fields (orchards)	24 ☐ industrial 25 ☐ recreational 26 ☐ religious 27 ☐ mechanical 28 ☐ technical 29 ☐ agricultural 30 ☐ commercial
6	34 ☐ ruins (incomplete buildings)	35 ☐ mesa (plateau)	36 ☐ waterfall 37 ☐ glacier 38 ☐ canal (channel, manmade waterway)	39 ☐ desert	40 ☐ forest 41 ☐ jungle 42 ☐ swamp (marsh)	31 ☐ wilderness 32 ☐ urban 33 ☐ rural (pastoral) 131 ☐ historical (archaeological)
5	43 ☐ isolated settlement 44 ☐ town (village) 45 ☐ city	46 ☐ single peak 47 ☐ hills (slopes, bumps, humps, mounds) 48 ☐ mountains 49 ☐ cliff(s) 50 ☐ [plain, delta] 51 ☐ valley (cleft, gully, irreg. depression) 52 ☐ canyon 53 ☐ [crater, bowl-shape, regular depression]	54 ☐ unbounded large expanse of water (ocean, sea) 55 ☐ completely bounded expanse of water (lake, pool, pond) 56 ☐ partially bounded expanse of water (bay) 57 ☐ island 58 ☐ river (stream, creek) 59 ☐ coastline		60 ☐ vegetation (trees)	

Figure A2. Coding form.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Figure A3. Coding form.

ABSTRACT DESCRIPTOR LEVELS I

Experiment: ______
Trial: ______
Resp./Targ: ______
Coder: ______
Viewer: ______

QUALITIES

LEVEL	COLOR	OTHER VISUAL	IMPLIED TEXTURE	IMPLIED TEMPERATURE	IMPLIED MOVEMENT	AMBIENCE
4	61 ☐ yellow	71 ☐ shiny (reflective)	80 ☐ smooth	85 ☐ hot	89 ☐ flowing	91 ☐ congested (cluttered, dense, busy)
	62 ☐ orange	72 ☐ [gold]	81 ☐ fuzzy	86 ☐ cold (snow, ice)	90 ☐ other implied movement	92 ☐ serene (peaceful, unhurried, unfrenetic)
	63 ☐ red	73 ☐ [silver]	82 ☐ grainy (sandy, crumbly)	87 ☐ humid		93 ☐ closed in (claustrophobic)
	64 ☐ blue	74 ☐ [chrome]	83 ☐ rocky (ragged, rugged, jagged, rubbled, rough)	88 ☐ dry (arid)		94 ☐ open (spacious, vast, expansive)
	65 ☐ green	75 ☐ [copper]	84 ☐ striated			95 ☐ ordered (aligned)
	66 ☐ purple (pink)	76 ☐ obscured (fuzzy, dim, smoky)				96 ☐ disordered (jumbled, unaligned)
	67 ☐ brown (beige)	77 ☐ cloudy (foggy, misty)				
	68 ☐ black	78 ☐ old				
	69 ☐ white	79 ☐ weathered (eroded, incomplete)				
	70 ☐ grey					

ARCHETYPES

LEVEL	STRUCTURE	ELEVATION	INTERFACE	UNIQUENESS	AMBIENCE
3	97 ☐ building(s) (structure(s))	98 ☐ rise (vertical rise as well as slope)	100 ☐ light/dark areas (big swaths)	104 ☐ single (or central) predominant feature	106 ☐ manmade (or altered)
		99 ☐ flat	101 ☐ boundaries	105 ☐ odd (or surprising) juxtaposition of elements	107 ☐ natural
			102 ☐ land/water interface		
			103 ☐ land/sky interface (horizon)		

Figure A4. Coding form.

ABSTRACT DESCRIPTOR LEVELS II

Experiment: ______
Trial: ______
Resp./Targ: ______
Coder: ______
Viewer: ______

2-D & 3-D GEOMETRIES

LEVEL	RECTILINEAR FORMS	CURVILINEAR FORMS	MIXED FORMS	IRREGULAR FORMS	REPEAT MOTIF
2	108 ☐ rectangle (square, box)	112 ☐ circle (oval, sphere)	114 ☐ cylinder	117 ☐ irregular forms (irregular features)	118 ☐ repeat motif
	109 ☐ triangle (trapezoid, pyramid)	113 ☐ [torus]	115 ☐ cone		
	110 ☐ other polygonal (> 4 sides: hexagon, octagon, etc.)		116 ☐ semicircle (hemisphere, dome)		
	111 ☐ cross-hatch (grid)				

1-D GEOMETRY

LEVEL			
1	119 ☐ stepped	127 ☐ arc (curve)	130 ☐ meandering curve
	120 ☐ parallel lines	128 ☐ wave form (ripples)	
	121 ☐ vertical lines	129 ☐ spiral	
	122 ☐ horizontal lines		
	123 ☐ diagonal lines		
	124 ☐ V-shape		
	125 ☐ inverted V-shape		
	126 ☐ other angles		

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Appendix B
Fuzzy Set Analysis Testbed

The following pages show the targets, responses, and analysis for two remote-viewing trials.

Figure B1. Target for Session 9004.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

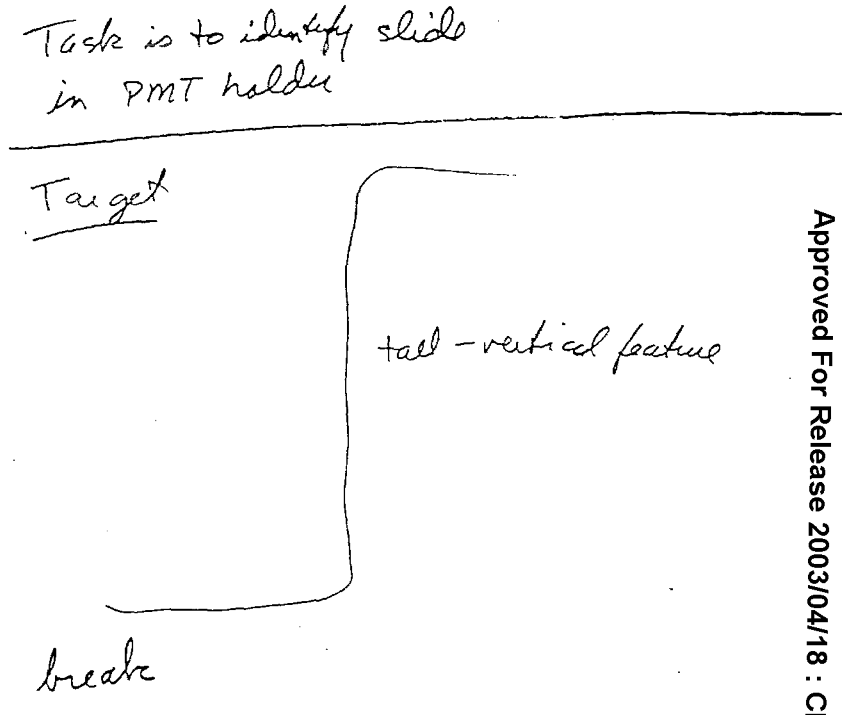


Figure B2. Page one of the response (Session 9004, Target 1094)

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

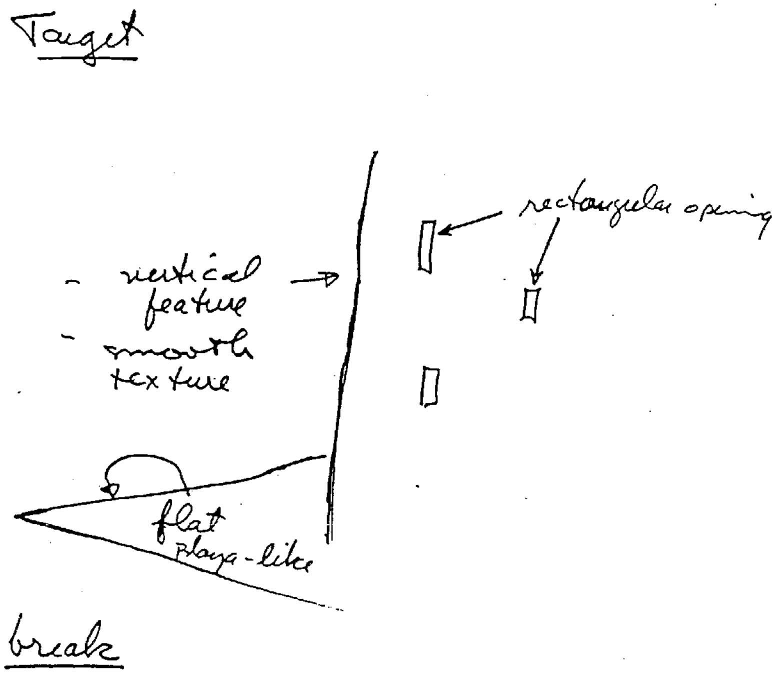


Figure B3. Page two of the response (Session 9004, Target 1094).

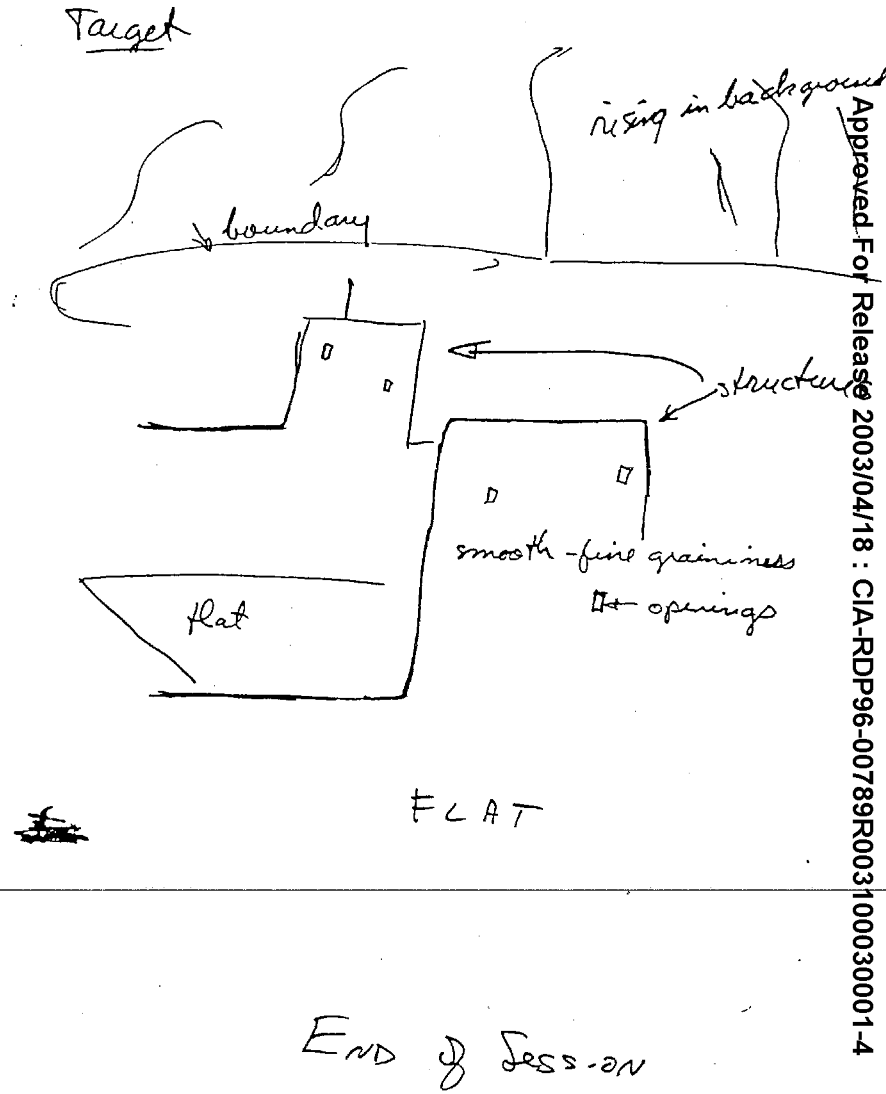


Figure B4. Page three of the response (Session 9004, Target 1094).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE B1
TARGET–RESPONSE 9004

Element	Name	Target	Response
20	Roads	0.30	0.00
23	Agricultural fields	0.05	0.00
32	Urban	0.00	0.50
33	Rural, pastoral	0.60	0.50
44	Town, village	0.00	0.50
45	City	0.00	0.40
46	Single peak	0.70	0.00
47	Hills, slopes, bumps, mounds	0.10	0.40
48	Mountains	0.00	0.60
49	Cliffs	0.00	0.10
60	Vegetation, trees	0.30	0.00
64	Blue	0.50	0.00
65	Green	0.30	0.00
69	White	0.10	0.00
70	Grey	0.20	0.00
76	Obscured, fuzzy, dim, smoky	0.20	0.00
77	Cloudy, foggy, misty	0.20	0.00
79	Weathered, eroded, incomplete	0.00	0.10
80	Smooth	0.00	1.00
81	Fuzzy	0.20	0.00
82	Grainy, sandy, crumbly	0.20	1.00
90	Other implied movement	0.20	0.00
91	Congested, cluttered, busy	0.10	0.30
92	Serene, peaceful, unhurried	0.40	0.00
93	Closed in, claustrophobic	0.00	0.10
94	Open, spacious, vast	0.60	0.00
95	Ordered, aligned	0.00	0.40
97	Buildings, structures	0.00	1.00
98	Rise, vertical rise, slope	0.60	1.00
99	Flat	0.30	1.00
100	Light/dark areas	0.10	0.00
101	Boundaries	0.30	1.00
103	Land/sky interface	0.50	0.00
104	Single predominant feature	0.60	0.00
105	Odd juxtaposition, surprising	0.30	0.00
106	Manmade, altered	0.20	0.80
107	Natural	0.70	0.20
108	Rectangle, square, box	0.00	1.00
109	Triangle, pyramid, trapezoid	0.60	0.00
115	Cone	0.60	0.00
117	Irregular forms	0.00	0.20
118	Repeat motif	0.10	0.60
119	Stepped	0.10	0.70
120	Parallel lines	0.10	0.00
121	Vertical lines	0.10	1.00
122	Horizontal lines	0.10	0.00
123	Diagonal lines	0.40	0.00
125	Inverted V-shape	0.70	0.00
126	Other angles	0.00	0.10

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

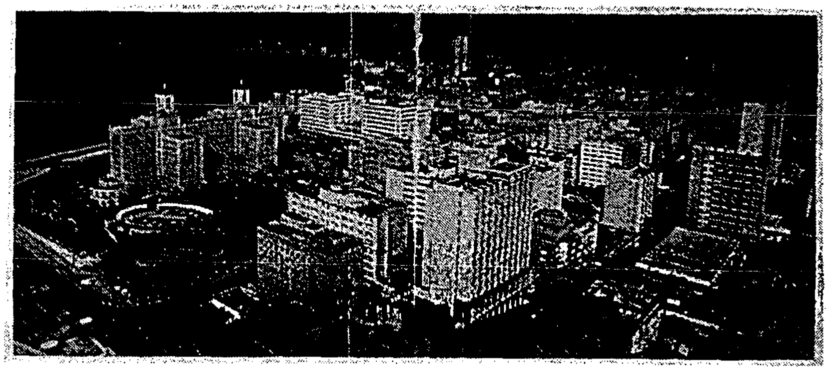


Figure B5. Target for Session 9005.

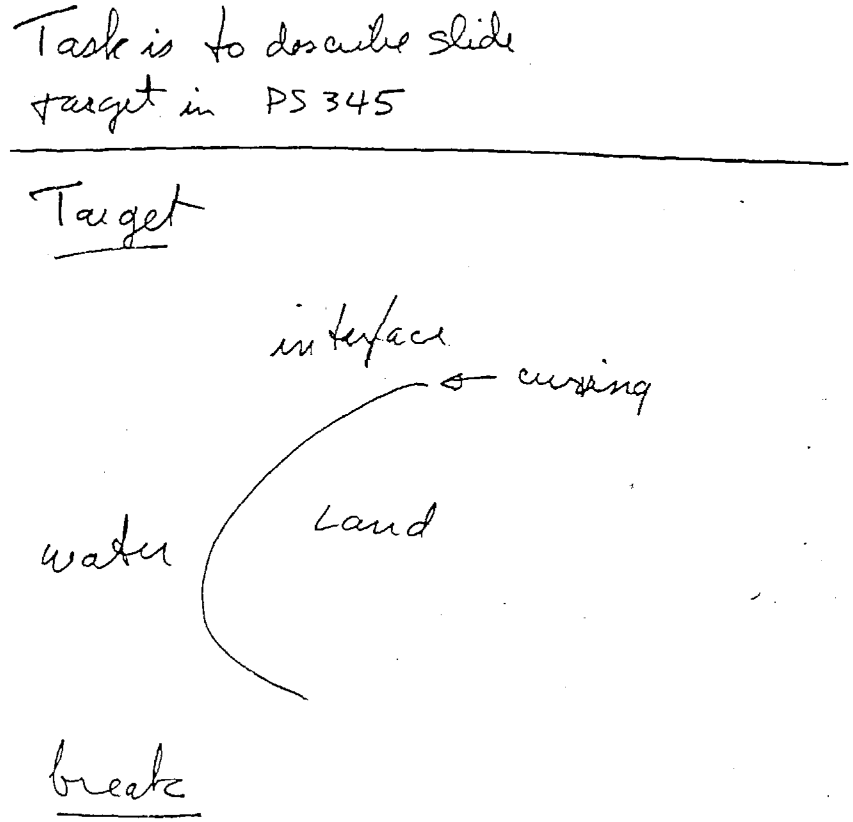


Figure B6. Page one of response (Session 9005, Target 1005).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

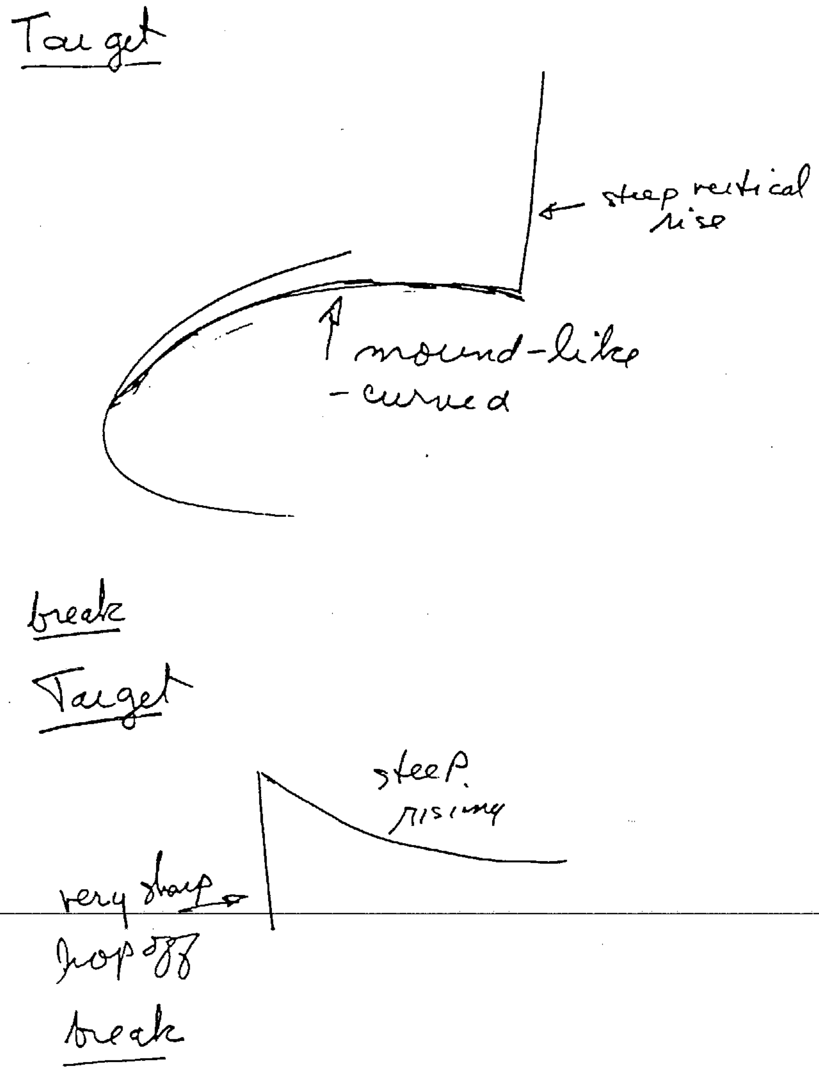


Figure B7. Page two of response (Session 9005, Target 1005).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

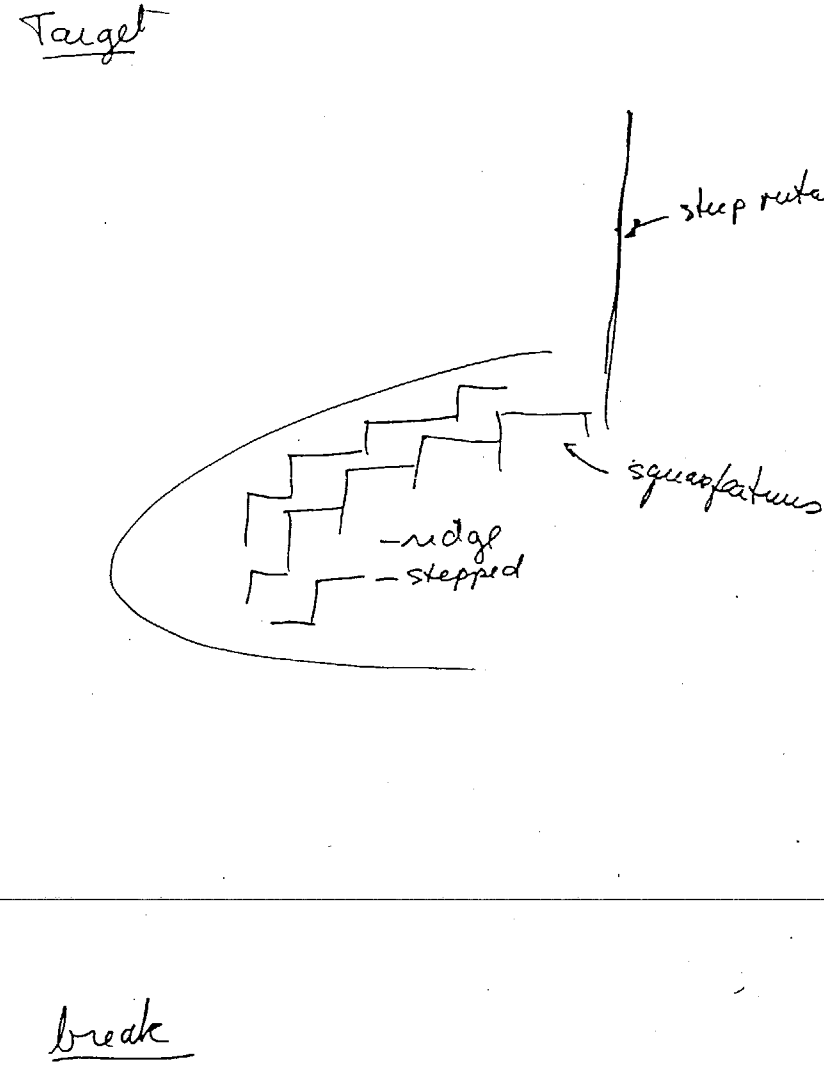


Figure B8. Page three of response (Session 9005, Target 1005).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

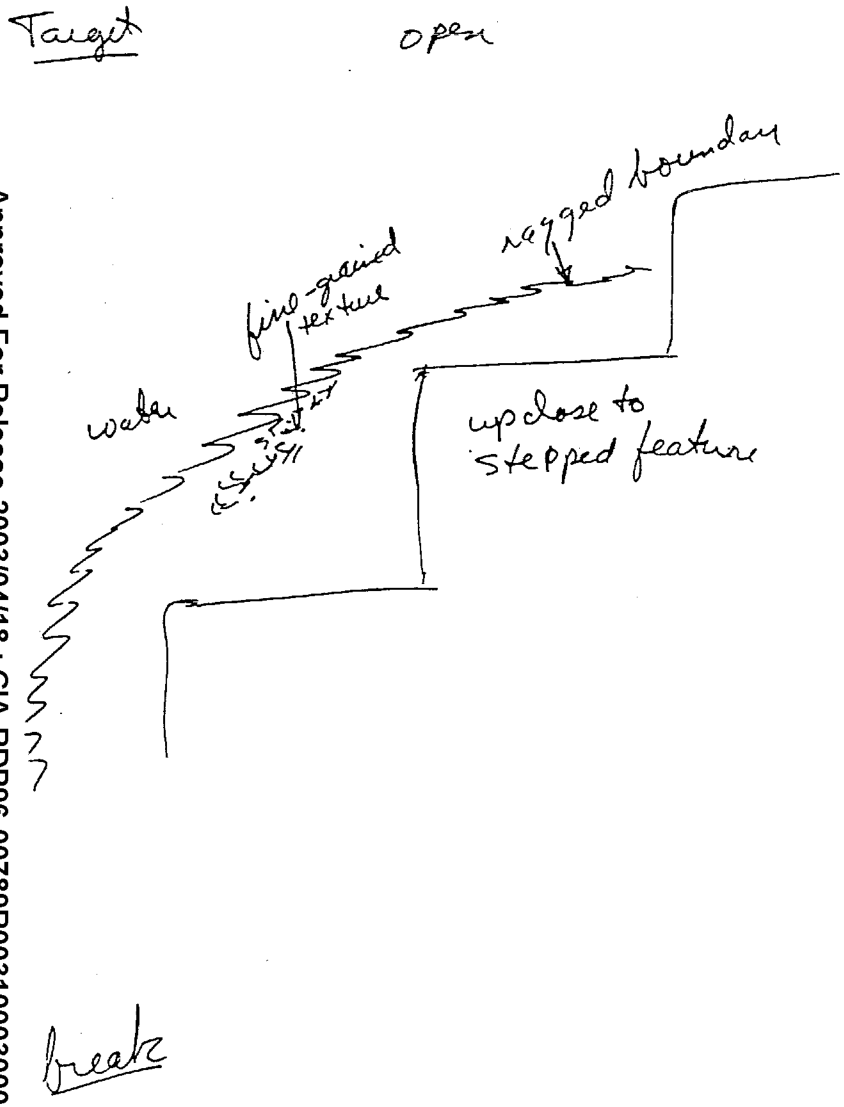


Figure B9. Page four of response (Session 9005, Target 1005).

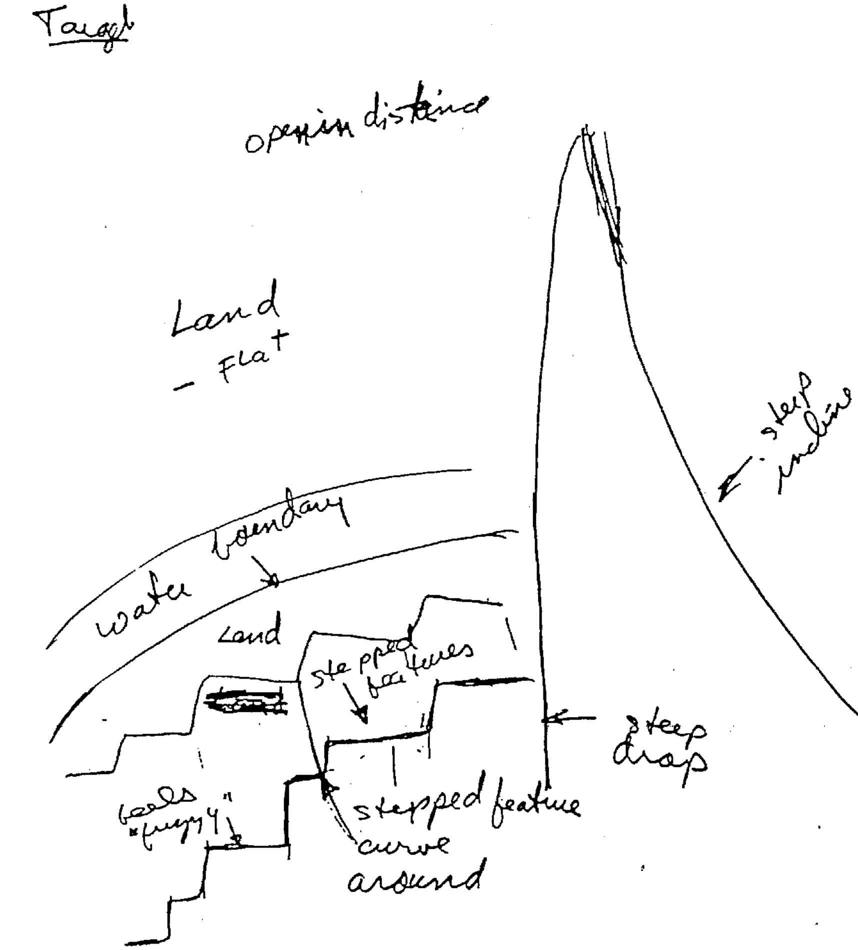


Figure B10. Page five of response (Session 9005, Target 1005).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table B2
Target–Response 9005

Element	Name	Target	Response	$T_{\alpha=0.2}$	T∩R
14	Spire, minaret, tower	0.00	0.20	0	0.00
20	Roads	0.10	0.10	0	0.00
32	Urban	0.80	0.70	1	0.70
38	Canal, manmade waterway	0.00	0.10	0	0.00
44	Town, village	0.00	0.30	0	0.00
45	City	0.90	0.70	1	0.70
46	Single peak	0.00	0.20	0	0.00
47	Hills, slopes, bumps, mounds	0.00	0.10	0	0.00
54	Unbounded large expanse water	0.00	0.40	0	0.00
56	Partially bounded water	0.30	0.30	1	0.30
58	River, stream, creek	0.00	0.40	0	0.00
59	Coastline	0.00	0.20	0	0.00
60	Vegetation, trees	0.20	0.20	1	0.20
64	Blue	0.25	0.00	1	0.00
65	Green	0.20	0.00	1	0.00
67	Brown, beige	0.50	0.00	1	0.00
69	White	0.10	0.00	0	0.00
70	Grey	0.10	0.00	0	0.00
80	Smooth	0.10	0.00	0	0.00
81	Fuzzy	0.00	1.00	0	0.00
82	Grainy, sandy, crumbly	0.00	1.00	0	0.00
83	Rocky, ragged, rubbled, rough	0.00	1.00	0	0.00
91	Congested, cluttered, busy	0.70	0.70	1	0.70
94	Open, spacious, vast	0.10	1.00	0	0.00
95	Ordered, aligned	0.00	0.30	0	0.00
96	Disordered, jumbled, unaligned	0.30	0.00	1	0.00
97	Buildings, structures	0.80	0.90	1	0.90
98	Rise, vertical rise, slope	0.00	1.00	0	0.00
99	Flat	0.50	1.00	1	1.00
100	Light/dark areas	0.10	0.00	0	0.00
101	Boundaries	0.20	1.00	1	1.00
102	Land/water interface	0.30	1.00	1	1.00
103	Land/sky interface	0.10	0.10	0	0.00
104	Single predominant feature	0.10	0.40	0	0.00
106	Manmade, altered	0.80	0.80	1	0.80
107	Natural	0.20	0.20	1	0.20
108	Rectangle, square, box	0.70	1.00	1	1.00
111	Cross-hatch, grid	0.30	0.00	1	0.00
112	Circle, oval, sphere	0.10	0.00	0	0.00
116	Semicircle, dome, hemisphere	0.10	0.30	0	0.00
118	Repeat motif	0.40	0.80	1	0.80
119	Stepped	0.20	1.00	1	1.00
120	Parallel lines	0.30	0.30	1	0.30
121	Vertical lines	0.50	1.00	1	1.00
122	Horizontal lines	0.10	0.00	0	0.00
123	Diagonal lines	0.10	0.20	0	0.00
125	Inverted V-shape	0.00	0.20	0	0.00
127	Arc, curve	0.30	1.00	1	1.00
128	Wave form	0.00	0.10	0	0.00
Totals			21.20	22.00	12.60

Accuracy = 0.573
Reliability = 0.594
Figure of merit = 0.340

Appendix C
"Ground Truth" Instruction and Coding Form

Analysts' Instructions for Remote-Viewing Series 900X

Thank you for helping us perform a *post hoc* assessment of a series of remote viewings. The targets were actually 35-mm slides that were attached to a photomultiplier, a device to measure small amounts of light. We were searching for possible physical correlates to remote viewing.

You will find in your packet 6 remote viewing responses labeled 9001–9006 respectively. Also shown is the target number of the intended photograph. We have supplied the original, rather than the 35-mm slide.

We would like you to make a *subjective* judgment as to the degree of correspondence between the remote viewing response and its associated target. Familiarize yourself with the task by first looking at all the responses and their intended targets. Then, on a session-by-session basis, rate your assessments. You are completely free to define what is meant by "Degree of Correspondence." Indicate your judgment by marking one line across the appropriate continuous scale shown below. A vertical line near the "None" end of the scale will indicate that you feel there is very little correspondence between that response-target pair. Likewise a vertical line near the "Complete" end of the scale will indicate that you feel that there is a significant degree of correspondence.

Many of the responses begin with a little information and build toward a composite drawing at the end. Please assess the response in its entirety as best you can. Thank you again.

SESSION	DEGREE OF CORRESPONDENCE (None ——— Complete)	TARGET
9001	|————————————|	1034
9002	|————————————|	1042
9003	|————————————|	1065
9004	|————————————|	1094
9005	|————————————|	1005
9006	|————————————|	1024

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

References

Dawes, R. M. (1988). *Rational choice in an uncertain world.* New York: Harcourt, Brace, Jovanovich.

Honorton, C. (1975). Objective determination of information rate in psi tasks with pictorial stimuli. *Journal of the American Society for Psychical Research,* **69,** 353–359.

Hubbard, G. S., May, E. C., & Frivold, T. J. (1987). Possible photon production during a remote viewing task: A replication experiment. Final Report, SRI Project 1291, SRI International, Menlo Park, California.

Humphrey, B. S., May, E. C., Trask, V. V., & Thomson, M. J. (1986). Remote viewing evaluation techniques. Final Report, SRI Project 1291, SRI International, Menlo Park, California.

Humphrey, B. S., May, E. C., & Utts, J. M. (1988). Fuzzy set technology in the analysis of remote viewing. *Proceedings of the 31st Annual Convention of the Parapsychological Association* (pp. 378–394).

Jahn, R. G., Dunne, B. J., & Jahn, E. G. (1980). Analytical judging procedure for remote perception experiments. *Journal of Parapsychology,* **44,** 207–231.

May, E. C. (1983). A remote viewing evaluation protocol. Final Report (revised), SRI Project 4028, SRI International, Menlo Park, California.

May, E. C., Humphrey, B. S., & Mathews, C. (1985). A figure of merit analysis for free-response material. *Proceedings of the 28th Annual Convention of the Parapsychological Association,* (pp. 343–354).

Puthoff, H. E., & Targ, R. (1976). A perceptual channel for information transfer over kilometer distances: Historical perspective and recent research. *Proceedings of the IEEE,* **64**(3), 329–354.

Solfvin, G. F., Kelly, E. F., & Burdick, D. S. (1978). Some new methods of analysis for preferential-ranking data. *Journal of the American Society for Psychical Research,* **72**(2), 93–110.

Targ, R., Puthoff, H. E., & May, E. C. (1977). State of the art in remote viewing studies at SRI. *1977 Proceedings of the International Conference of Cybernetics and Society* (pp. 519–529).

Zick, R., Carlstein, E., & Budescu, D. V. (1987). Measures of similarity among fuzzy concepts: A comparative analysis. *International Journal of Approximate Reasoning,* **1**(2), 221–242.

SRI International
333 Ravenswood Av.
Menlo Park, CA 94025

Division of Statistics
University of California, Davis
Davis, CA 95616

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

OBSERVATION OF NEUROMAGNETIC FIELDS IN RESPONSE TO REMOTE STIMULI

by

Edwin C. May, Wanda W. Luke, Virginia V. Trask, and Thane J. Frivold
SRI International, Menlo Park, California

ABSTRACT

We have conducted a conceptual replication of an SRI/Langley Porter study in which a single subject's central nervous system (CNS) responded to a remote, and isolated flashing light. The CNS activity of eight remote viewers was monitored by a seven-channel magnetoencephalograph (MEG). Visual stimuli were randomly presented to an isolated individual who acted as a "sender" while MEG data were collected from a viewer (receiver). The stimuli were 5-cm square, linear, vertical, sinusoidal gratings lasting 100 ms (remote stimuli). Time markers were randomly inserted into the data stream as control points (pseudo stimuli). The dependent variable was the root-mean-square (RMS) average phase shift of the dominant alpha frequency. Using a Monte Carlo technique to estimate p-values, we observed significant (combined across all viewers) RMS phase shifts resulting from the remote stimuli ($Z_s = 1.99, p \leq 0.024$, *effect size* $= 0.599$). Similarly, the combined statistic for the pseudo stimuli was also significant ($Z_s = 2.92, p \leq 0.002$, *effect size* $= 0.924$). The phase shifts from the remote and the pseudo stimuli are independently *not* characteristic of the data at large. This result was unexpected, and suggests that we may have observed a CNS response to an unintended stimulus (i.e., electromagnetic interference, EMI, from the computing hardware). However, in the SRI/Langley Porter study, EMI had been eliminated, thus, it remains possible that the CNS changes resulted from an anomalous form of information transfer.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

I INTRODUCTION

1. Physiological Correlates to Psychoenergetic Functioning: A Brief History

Evidence from several laboratories has indicated the possible existence of an as-yet-unidentified channel wherein information is coupled from remote electromagnetic stimuli to the human nervous system. Usually, the coupling has been indicated by physiological responses, even though there was no evidence of cognitive awareness of these stimuli. Physiological measures have included a plethysmographic response[1*] and electroencephalogram (EEG) activity.[2,3] Kamiya, Lindsley, Pribram, Silverman, Walter, and others have suggested that the whole range of EEG activity, including evoked potentials, spontaneous EEG, and the contingent negative variation (CNV) might be sensitive indicators of responses to any remote stimuli.[4]

In 1974, SRI International conducted a pilot study that investigated a single remote viewer's central nervous system (CNS) response to a remote light stimulus.[5] In this experiment, the viewer was asked to focus attention on a remote flashing (16-hertz [Hz]) light. Control periods (no light flashing) were randomly mixed with effort periods (light flashing). The viewer was further asked to register when he[†] perceived the flashing light by pressing a button.

During this pilot experiment, the viewer showed a significant[‡] decrease in alpha production when the remote light was flashing, compared with when the light was off. His button presses were random, however, indicating he was not cognitively aware of the flashing light. Two replications of this experiment were conducted with the same viewer at Langley Porter Neuropsychiatric Institute in San Francisco by Drs. David Galin and Robert Ornstein.[6] In the first of two experiments, the viewer continued to show a significant decrease of occipital alpha production only under the remote flashing light condition. In a second experiment conducted 3 months later, however, the viewer demonstrated a significant increas[e in] occipital alpha production.

Although we found that significant correlati[ons] appear to exist between the times of light flas[hes] and CNS activity, we considered this result to [be] only suggestive, with a definitive conclusion [re]quiring further experimentation.

With the advent of more sensitive CNS monitori[ng] equipment, known as magnetoencephalograp[hy] (MEG), and with an additional 15 years of remo[te] viewing experience, SRI conducted an experime[nt] to explore possible correlations between CNS acti[v]ity and remote stimuli. This experiment is the su[b]ject of this report.

2. Technological Background

Magnetoencephalography is a noninvasive technique used to measure, in three-dimensional space magnetic fields produced by neuronal electric currents in the cortex of the brain. A magnetoencephalography device (MEG) can determine the spatial distributions of specific groups of neurons participating in a given activity and their patterns of activity over time. This technology has been used in research ranging from evaluating how normal brains process information to diagnosing clinical conditions such as epilepsy and dementias.[7]

Neurons that participate in a given functional activity communicate between themselves and ultimately other parts of the body by a complex combination of electrical signals and chemical interactions. It is beyond the scope of this report to describe the cellular physiology involved, but is sufficient to say that this activity produces magnetic fields (predominantly dipole) that can be sensed externally.

The sensing device of a MEG is a cryogenic superconducting quantum interference device (SQUID) coupled with a gradiometer. SQUIDs currently being used are cooled by liquid helium. At a few degrees above absolute zero, an electrical current can flow through a superconductor with no applied voltage. The material of the SQUID consists of superconducting loops with two sections of thin insulating material connecting them (Josephson Junctions). This configuration is referred to as a DC SQUID. Some electrons can tunnel through this insulation. The

* References are at the end of this report.

† To keep the identity of the viewers confidential, we use the pronouns *he* and *his* throughout this report, regardless of the viewer's gender.

‡ Throughout this report, the word "significant" conforms to the standard definition; $p \leq 0.05$.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

presence of a weak magnetic field produces a phase difference for the wave function of the magnetic field [and] produces a phase difference for the wave function of the electrons across this barrier. The resulting interference pattern produced by the two different wave functions on each side of the barrier can be used to indicate the strength of these extremely weak magnetic fields.

The neuronal magnetic fields from the human brain are only about 10^{-13} tesla, while the earth's magnetic field is 10^{-4} tesla and normal urban noise is about 10^{-7} tesla. Care must be taken, therefore, to assure that the signal-to-noise ratio is favorable. This has been taken into consideration by the manufacturer of MEG equipment (BTi of San Diego, California), who has designed highly shielded sensors that use a second-order coupled gradiometer to reduce the environmental noise by about 10^6. The use of an aluminum and μ-metal magnetically shielded room can further reduce the noise by a factor of 10^3. If used together, these two precautionary measures can reduce the ambient noise by a factor of about 10^9—equivalent to the internal SQUID noise.

Because a MEG responds best to neuronal currents that are parallel to the skull (i.e., currents producing magnetic fields oriented tangentially to the skull), neuronal currents perpendicular to the skull may be missed. In reality, however, few neuronal electrical currents are exactly perpendicular to the skull, so some tangential component is almost always available to the SQUID.

Searching for a closely packed group of neurons can be a slow and tedious process. Due to technological restraints, a maximum of seven sensors can be used simultaneously to gather MEG measurements. Sensors on a seven-channel MEG are located on a 2-cm equilateral triangular grid forming the center and vertices of a regular hexagon. A subject wears a spandex cap with grid marks lined up with his nasion, inion, and earlobes to serve as a head-centered coordinate system. To identify the location of a neuronal-equivalent current dipole, many measurements have to be taken. Isocontour maps of field strength are used to represent the amplitude and polarity distribution of the magnetic fields. A least-squares procedure is applied to the observed fields to estimate the location of neuronal sources and orientation of the equivalent current dipole.[8] The estimated location of the neuronal source can then be identified anatomically with a magnetic resonance image scan of the head. Developments in technology may soon allow for enough channels to cover the whole head at once, thereby reducing data collection time and increasing precision.

MEG technology is based on a cryogenic SQUID operating in liquid helium. Because the Dewar flask cannot exceed a 45-degree angle, subjects must lie prone beneath the apparatus. MEG sensors are not attached to the head, but are lowered into position over the skull; the subject cannot move his head during monitoring without disturbing the measurement. For these two reasons, MEG equipment is not suited for long-term monitoring of a subject. These problems may be solved in the near future as new technology, such as high-temperature SQUIDs, develops.

A response from the MEG is a complex waveform consisting of a series of negative and positive peaks or components. Specific components of this waveform can be correlated with perceptual and cognitive processes. The most commonly observed response to a visual or auditory stimulus, for example, is a large component occurring approximately 100 ms after the onset of the stimulus. One hundred milliseconds appears to be the average latency period between stimulus and the first correlated neuronal activation in the brain.[8]

The earlier EEG technology measures electric potential, or event-related potentials (ERPs) produced by the electrical activity of the brain. A MEG measures the magnetic fields, or event-related fields (ERFs) produced by the electrical activity of specific groups of active neurons in the cortex. An EEG and a MEG, therefore, reveal different aspects of the electrical activity of the brain and are often used as complementary technologies. In some areas, however, the MEG technique has definite advantages over the EEG:

(1) ERPs taken from the scalp provide little information regarding the precise three-dimensional distribution of the neuronal sites producing the electrical activity. Brain tissues of unknown electrical conductivity and thickness, individual variations in skull thickness and geometry, and proximity to openings in the skull all make obtaining such detailed information difficult. The same is not true when using a MEG. Neuronal magnetic fields can travel through brain tissues without being significantly altered; this property, coupled with the dipole model, results in high spatial resolution of the neuronal activity.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

(2) EEG procedures are occasionally costly and can be invasive: EEG electrodes must be attached directly to the skull or to the brain of the subject, whereas MEG sensors are extracranial and are simply lowered into position against the skull.

(3) There is much controversy over the appropriate reference electrode in EEG work (a reference electrode is required with electric potential measurements, because only differences in electric potential are measured). There is no such problem with a MEG, because the measurement of magnetic fields is absolute.

II METHODS OF APPROACH

Our goal was to conduct a conceptual replication of the earlier SRI/Langley Porter experiments. Our basic hypothesis is that a viewer's CNS would respond to a remote light stimulus.

1. General Description

Using a seven-sensor MEG in a shielded room, we investigated the occipital-cortex neuronal magnetic activity that might occur in response to a remote "visual" stimulus.

The following definitions may be helpful:

- Viewer—An individual who attempts extrasensorimotor communication with the environment (e.g., the perception of remote stimuli).
- Direct Stimuli (DS)—Visual stimuli occurring within the normal visual sensory channels.
- Sender—An individual who, while receiving direct stimuli, acts as a putative transmitter to a remote individual (i.e., viewer) who is attempting to receive the same information via extrasensorimotor communication.
- Remote Stimuli (RS)—Visual stimuli occurring outside the normal range of known sensory channels.
- Pseudo Stimuli (PS)—A time marker in the data stream with no associated stimuli.

In this report, a direct stimulus to the sender is also considered as a remote stimulus to the viewer.

2. Protocol

2.1 General Considerations

To begin a session, a sender is isolated in a room while a viewer is monitored by a MEG in a shielded room about 40 m away. Only the sender is presented with a number of direct visual stimuli at random intervals within a 120-second period, the length of one run. One session usually consists of 10 runs.

2.1.1 Viewers

Eight viewers were selected for this experiment. Four were known to be good remote viewers, and four were staff members with unknown viewing ability. Each viewer contributed a minimum of one and a maximum of three independent sessions.

2.1.2 Senders

The senders in all sessions were either various staff members who were well known to the viewers or they were spouses.

2.1.3 Dependent Variable

The dependent variable is the root-mean-square (RMS) phase shift of the primary alpha activity as a result averaged over all RS.

2.2 Specific Protocol Details

2.2.1 Stimuli

Remote stimuli consisted of a standard video encoded blank screen with a 5-cm square, linear, vertical, sinusoidal grating lasting about 100 ms. These stimuli (DS to the sender) subtended 2 degrees in the lower left visual field of the sender. This was maintained by asking the sender to focus his visual attention on a permanent mark on the monitor. During the experiments described in this report, no attempt was made to monitor the sender in any way. Pseudo stimuli consisted of the blank screen without the superimposed grating, and were included as a putative within-run control.

2.2.2 Run Timing

Figure 1 shows a schematic timing diagram for one run. No two stimuli of any type were allowed to occur within a 3-second period of each other. A stimulus may occur, however, any time within a

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

4.5-second window thereafter. The sender was presented with a minimum of 9 and a maximum of 15 DS occurring at random intervals within a 120-second period. In all but the first session, a random number of pseudo stimuli (i.e., random time markers with no concomitant stimuli—PS) were added as a within-run control. A viewer was never presented with direct stimuli except in locating the maximal response to the visual areas (see Section II.2.2.4).

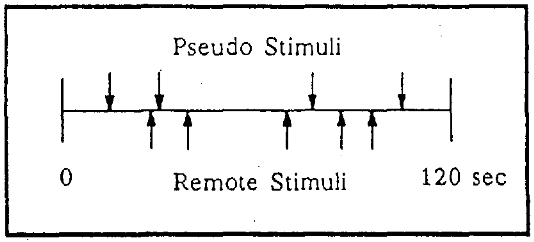


Figure 1 Schematic Timing Protocol—Single Run

2.2.3 Instructions to Viewers

In all sessions, the viewers were completely informed about the details of the experiments. Prior to their placement on the MEG table, they were shown the location of the RS display monitor, and were instructed to place their attention upon it or the sender during the session.

For some sessions, the viewer was instructed to press a fiber-optic-coupled button when he felt that he perceived stimuli. Each button press was marked in the data record. Button pressing was retained in this protocol as part of the conceptual replication.

2.2.4 Sensor-array Placement and Calibration

We selected the location for the sensor array by optimizing the viewer's response to direct visual stimuli. Inherent in this choice is an assumption that may not be valid: namely, that neurons participating in a reaction to RS are the same as those that respond to DS. The sensor locations were then marked on an acetate transparency to allow for accurate repositioning of the sensors in later sessions. One such placement (right occipital—minus centimeters from the inion indicate the right hemisphere) is shown for viewer 002 in Figure 2. It should be noted that MEG sensor placements do not necessarily correspond to conventional EEG electrode placement.

For a calibration, the viewer was fitted with a spandex cap with grid marks aligned with his inion, nasion, and earlobes (i.e, head-centered coordinate system). The viewer was then placed as comfortably as possible on an observation table beneath the MEG. He must lie face down and look though a hole in the table to view the DS via a system of mirrors. These stimuli were displayed by a projector located outside the entrance to the shielded room. The sensors of the MEG were lowered from above to touch his head over the right occipital lobe. In this configuration, the sensor array was moved at the end of 30 DS to a position that optimized his response to the DS. Once found, the array position was marked on the cap for subsequent repositioning.

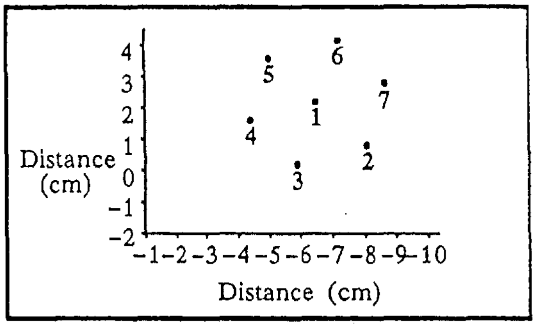


Figure 2 Sensor Position Relative to the Inion (0,0) for Viewer 002

2.2.5 Sequence of Events for a Session

The following is the schedule of events for a session:

- Collect approximately 10 minutes of background data with no viewer or sender present and the MEG in full operation.
- Isolate the sender with the stimulus display device.
- With the viewer on the table, position the sensor array at the calibration point.
- At time = 0, start the monitoring of data with computer-generated trigger. Data are collected the entire 120 seconds at a rate of 200 samples per second.
- At time < 120 seconds, present 9 to 15 remote and 9 to 15 PS to the sender.
- At time > 120 seconds, allow the viewer to relax for about 2 to 5 minutes without leaving the table. This break generally consists of the sender entering the shielded room to engage the viewer in conversation.
- Collect nine additional runs with the same procedure while the viewer remains positioned on the table under the MEG.

3. Data Analyses

If our initial assumption about sensor positioning is true, and if the earlier results are replicated, we expect to see a change in alpha production as a result of the RS. We might also expect an evoked response similar to visual ERFs. Figure 3 is an idealized illustration of these expected results in the time-series data. Times less than zero are prestimulus; times greater than zero are poststimulus. The stimulus lasts 100 ms.

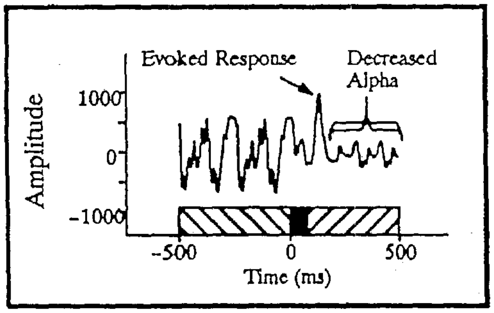


Figure 3 Idealized Results for a Single Stimulus

For each session, the following was computed for each RS and PS, respectively:

(1) Five hundred ms of pre- and post-stimulus time-series data were separately detrended and filtered (40 Hz lowpass).

(2) The power spectrum was computed for each 500-ms pre- and post-stimulus period.

(3) The relative phase change of the dominant alpha frequency from pre- to post-stimulus period was computed as the arctangent of the ratio of the imaginary and real component of the transfer function. The transfer function is defined as the ratio of the FFT of the post-stimulus period divided by the FFT of the pre-stimulus period.

(4) One thousand ms of time-series data (i.e., 500 ms pre- and post-stimulus) was separately detrended and filtered (40 Hz lowpass).

In addition, the following averages were computed across all RS and PS, respectively:

(5) The average power pre- and post-stimulus.

(6) The root-mean-square (RMS) average phase shift.

(7) The 1000-ms time average of the pre- and post-stimulus periods taken as a single record.

(8) The "power spectra" of the pre- and post-stimulus time averages were computed. (We recognize that a power spectrum of a time average is not an accurate representation of the average power spectrum, however it is an indicator of phase shift.)

4. Monte Carlo Calculations

The analysis of CNS activity has always been problematic, because alpha bursts lasting from 0.1 to a few seconds occur at random intervals. From a statistical point of view, the data fail to satisfy at least two underlying assumptions of the usual statistical methods (e.g., ANOVA and MANOVA). Most standard statistical tests assume that all samples of the data are independent. MANOVA can be configured to remove this particular assumption, nonetheless, it and the other tests assume that the process under study is stationary; that is, whatever the statistical properties are, they remain constant over time. In other words, the measured properties should not depend upon when the activity is sampled. CNS time series data do not satisfy either of these assumptions.

To avoid these difficulties, and to obtain probability estimates of the observed RMS phase shifts, we adopted a simple Monte Carlo approach. In the usual statistical analysis, the phase shift is compared to an ideal distribution, or its likelihood of occurrence is computed using some nonparametric technique. Both techniques attempt to determine the degree to which the observed phase shift is exceptional, given the universal set of all possible data. The Monte Carlo method that we used, however, can only determine the degree to which the observed phase shift is exceptional, given the available data sample. Thus, a new Monte Carlo estimate must be computed for each individual data set.

The general Monte Carlo procedure is as follows:

(1) Using the same timing algorithm to create the original RS, generate N sets of M stimuli, where M is the number of original RS.

(2) For each pass $(1...N)$, compute the RMS phase shift averaged over M remote stimuli.

(3) Sort the resulting N values to form the RMS phase shift distribution in the given data sample.

(4) Compute the probability that the observed value would be as large (or larger), given a repeated random sample of the data. Note that

this p-value is *not* the probability that the measure is as large, given a different data sample.

We have used this technique to compute p-values for the RMS phase shifts throughout this report.

III RESULTS

Eight viewers (002, 007, 009, 372, 374, 389, 454, and 531) from SRI International participated in the effort. Viewers 002, 009, 372, and 389 were experienced, with strong track records. Viewers 007, 374, and 531, had not previously participated in remote viewing experiments. Viewer 454 had participated in novice remote viewing training and has produced significant evidence of remote viewing ability.

1. Calculations

To illustrate the reduction of the raw data, we use the 25 September 1988 session from viewer 002.

Figure 4 shows the time average over all RS of the amplitude (femto Tesla) of the magnetic CNS activity of viewer 002's response to RS. The data from all seven sensors are displayed in a pattern that is similar to the physical sensor array. Each sensor is labeled in a highlighted box. The number of stimuli comprising the average (118) is shown in the key. The onset of the 100-ms stimulus is represented at *time* = *0*, so negative time represents the pre-stimulus period and positive time represents the post-stimulus period. The total time period shown is 1 second. Because the stimuli are at random times relative to any uncorrelated CNS activity, averaging has reduced random single-stimulus amplitudes by $\sqrt{n}$ where n is the number of stimuli. Sensor 7 shows a clear change from a slow, regular alpha rhythm during the pre-stimulus period, to one of higher frequency, post-stimulus.

Figure 5 shows this change of alpha in the frequency domain. For each sensor, the power spectrum of its corresponding time series is displayed from 0 to 40 Hz. The power spectra are shown independently for the pre- and post-stimulus periods (separated by a dashed vertical line). Sensor 7 shows a strong 10-Hz peak pre-stimulus that vanishes post-stimulus Similar alpha reductions can be seen in all of the other six sensors.

The power spectrum of a time series average is *not* an indicator of the average power spectrum of the CNS activity, because time averages are phase sensitive and power spectra are not. Figure 6 illustrates this by showing the average power spectra (i.e., calculated on a stimulus-by-stimulus basis and then averaged) for the pre- and post-stimulus periods. There was little change of CNS power across the stimulus boundary throughout the frequency range.

Because a time average is sensitive to relative phase and a power spectrum is not, these data suggest that a relative phase shift occurs between pre- and post-stimulus periods. Figure 7 shows this relative RMS phase shift computed from 0 to 40 Hz for all sensors. As was the case for the time-series data, the RMS average was computed over $n = 118$ RS. In accordance with the protocol (Section II.3), the dependent variable was the RMS phase only at the dominant α-frequency.

At this point we are unable to determine if the variations seen in Figures 4 through 7 are meaningful. Toward that end, the identical quantities for the PS are shown in Figures 8 through 11. The "power" of the time averages for the remote stimuli differ markedly from those of the PS spectra (Figures 5 and 9).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

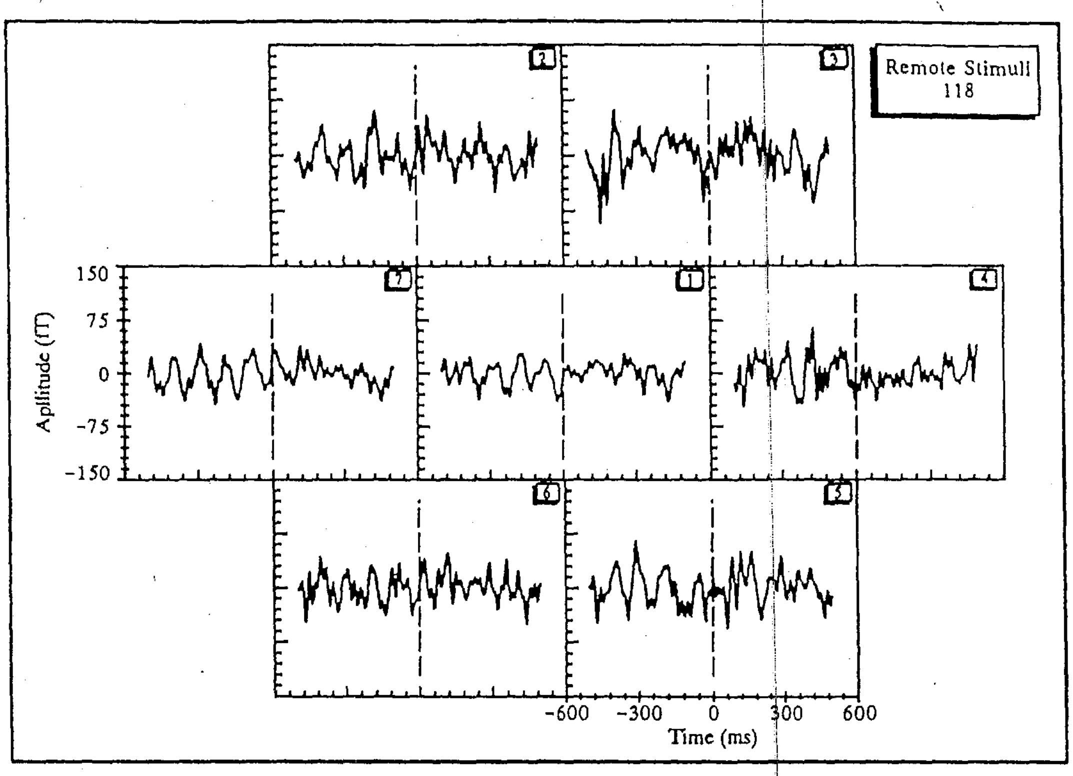


Figure 4 Viewer 2: Date 8/25/88: Session 1: Time Average

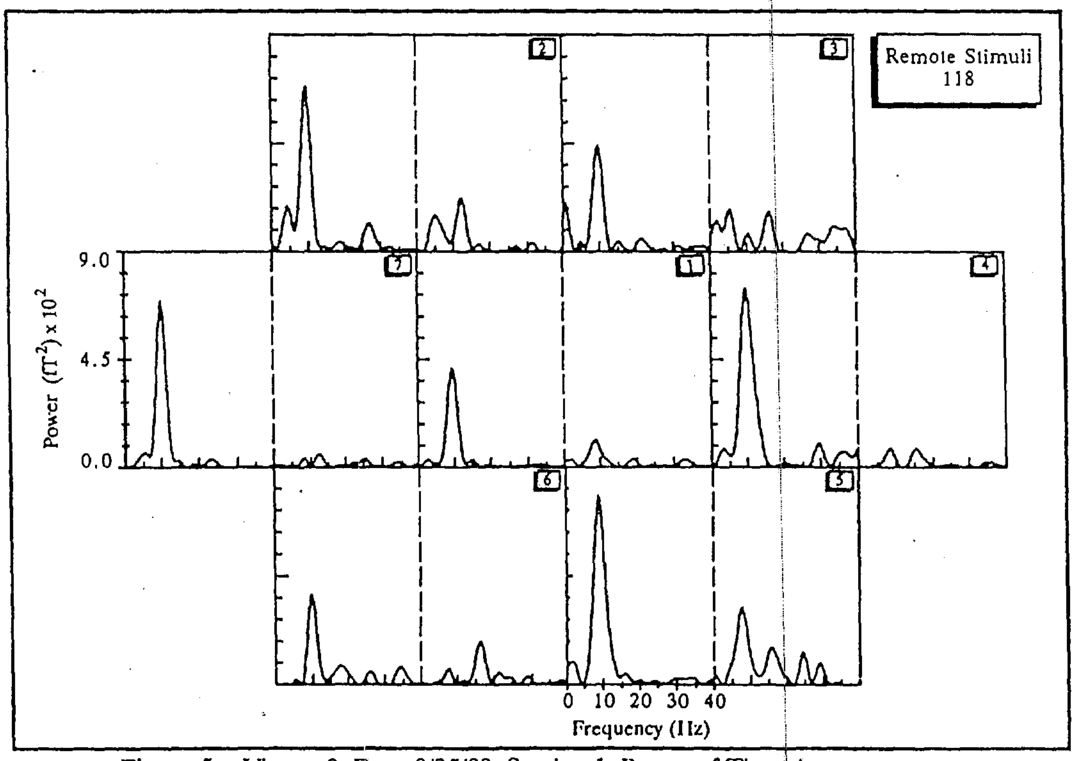


Figure 5 Viewer 2: Date 8/25/88: Session 1: Power of Time Average

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

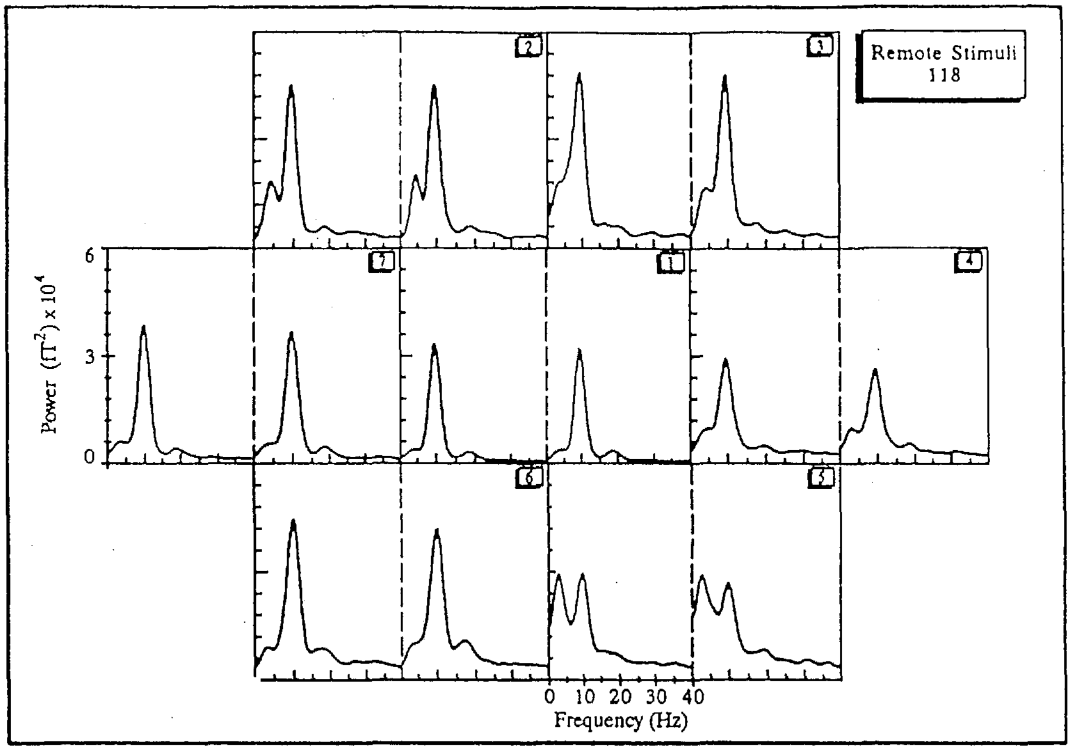


Figure 6 Viewer 2: Date 8/25/88: Session 1: Average Power

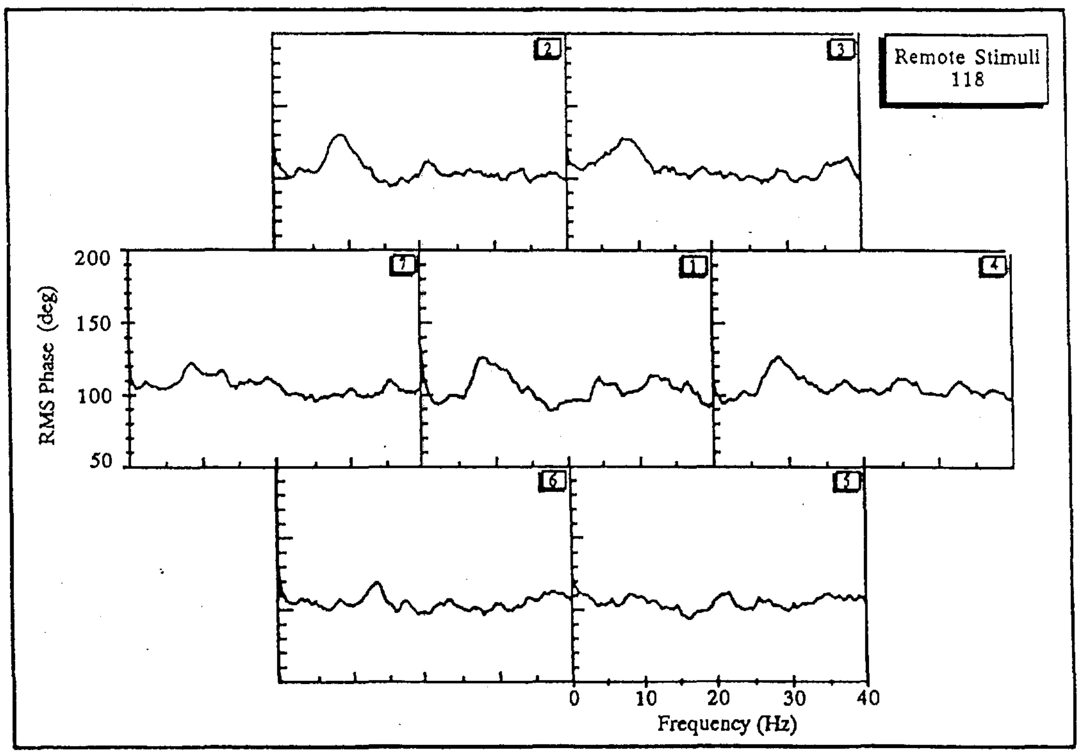


Figure 7 Viewer 2: Date 8/25/88: Session 1: RMS Phase

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

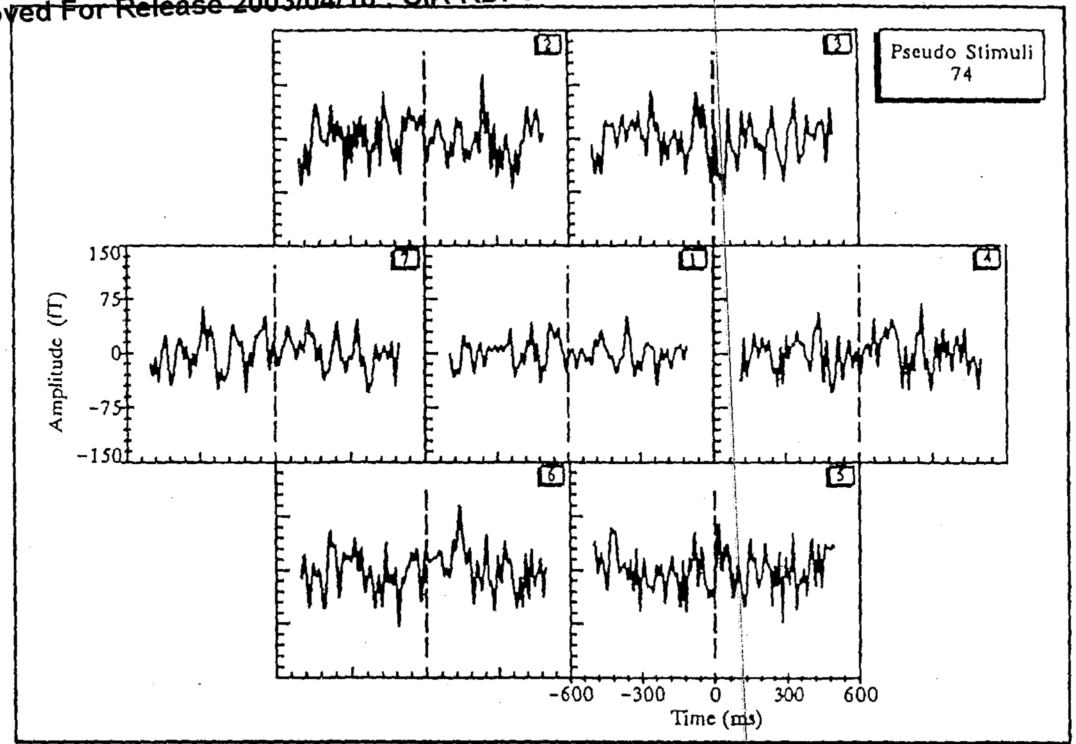


Figure 8 Viewer 2: Date 8/25/88: Session 1: Time Average

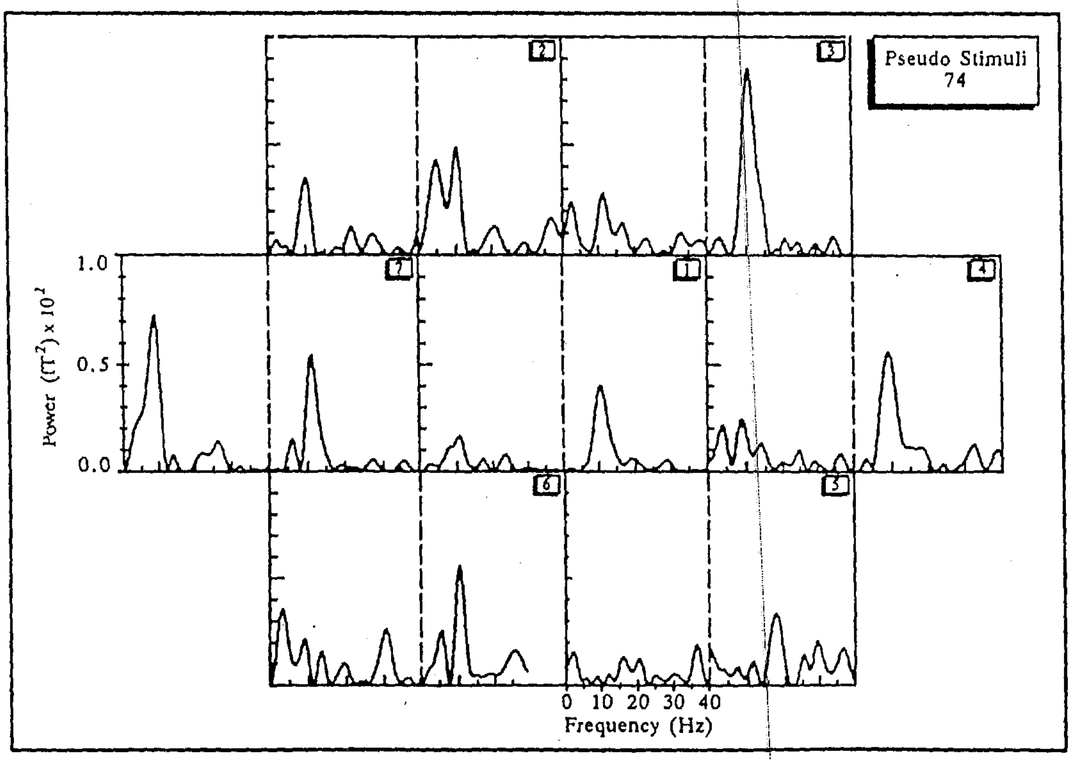


Figure 9 Viewer 2: Date 8/25/88: Session 1: Power of Time Average

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

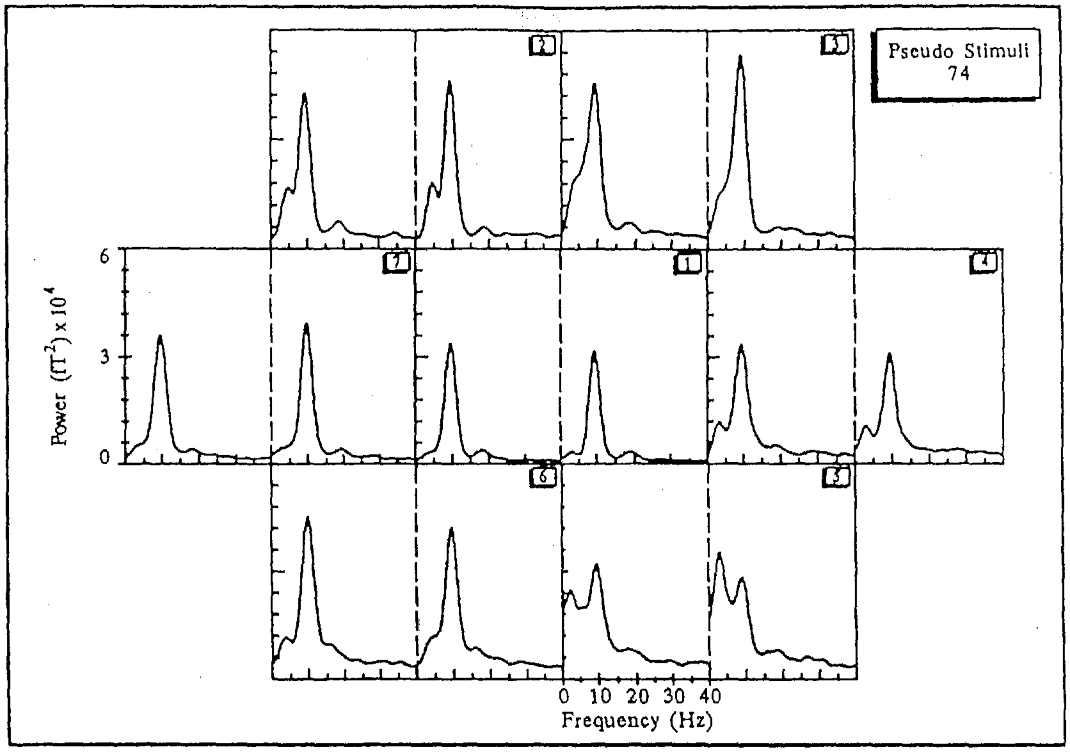


Figure 10 Viewer 2: Date 8/25/88: Session 1: Average Power

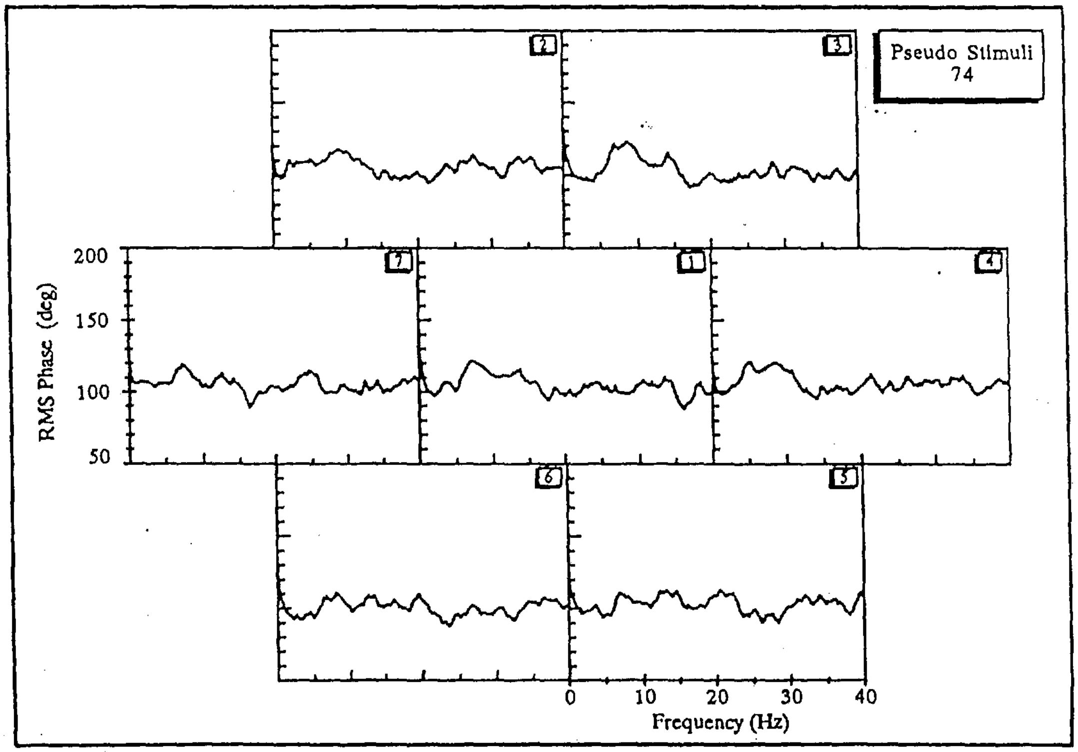


Figure 11 Viewer 2: Date 8/25/88: Session 1: RMS Phase

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

2. Monte Carlo Estimates of Significance

To determine if the changes that are seen qualitatively are exceptional, we analyzed the data by the Monte Carlo procedure outlined in Section II.4. We simulated the RS by generating 500 sets of Monte Carlo stimuli using the same random timing algorithm and number as in the original data. For each set, the RMS phase was calculated as described in Section II.3. The resulting 500 Monte Carlo RMS phases were sorted as a descending array, and the fraction of phases equal to or larger than the observed RS value was represented as a p-value. (The p-value is bounded on the low end by 1/500.) Figure 12 shows a histogram of one such Monte Carlo run, again using the data from viewer 002 as an example. The values of the RMS phase for the remote and pseudo stimuli are marked by vertical lines (see the key in Figure 12).

In accordance with the earlier study[6] in which we observed changes in alpha power, we established a single criterion for the selection of a sensor for analysis: the pre-stimulus average alpha power above background is larger than it is in any other sensor. Table 1 shows the viewer identification, date, sensor chosen for analysis, and the p-value (as defined above) for the RMS phase shift for the remote and pseudo stimuli, respectively.

The p-values shown in Table 1 are all single tailed (i.e., the area in the upper tail). Because the distribution of means is approximately normal, we have converted the empirical p-values to their respective two-tailed z-scores. If the p-value was less than 0.5, the z-score shown in Table 1 was computed from the inverse normal distribution assuming a p-value twice the one shown. If the p-value was more than 0.5, we subtracted it from 1.0, doubled the result, and computed the z-score as above. To test the null hypothesis that the combined RS phase shifts are characteristic of the data, we computed a standard Stouffer's Z (Z_s) for the 11 sessions shown in Table 1. There is statistical evidence that the data within $\pm$ 0.5 seconds of the RS are *not* characteristic of the data at large ($Z_s = 1.99, p \leq 0.024$, *effect size* $= 0.599$). Similarly, the combined statistic for the PS indicates that these data are also *not* characteristic ($Z_s = 2.92, p \leq 0.002$, *effect size* $= 0.924$). Therefore, there appears to be some statistical anomaly associated with the RMS phase shifts for both stimuli types.

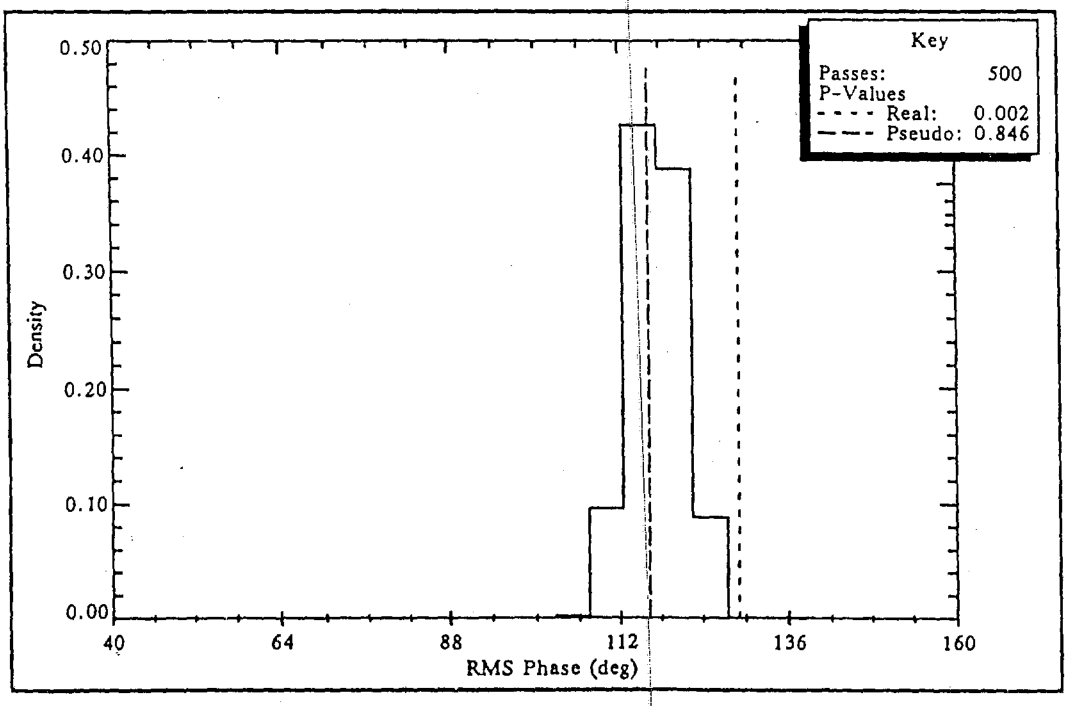


Figure 12 Viewer 2: Date 8/25/88: Session 1: RMS Phase: Sensor: 2: RS = 118

Table 1

Results of Monte Carlo Calculation for RMS Phase

I.D.	Date	Sensor	P-Value (1-tail)		Z-Score (2-tail)	
			Remote	Pseudo	Remote	Pseudo
009	06/24/88	6	0.650	-	-0.524	-
002	08/25/88	2	0.002	0.848	2.653	0.513
	08/26/88	6	0.904	0.966	0.871	1.491
372	10/19/88	7	0.094	0.168	0.885	0.423
374	03/29/89	6	0.154	0.810	0.501	0.305
007	03/29/89	7	0.970	0.180	1.555	0.358
389	05/23/89	4	0.288	0.040	-0.191	1.405
	05/24/89	5	0.260	0.016	-0.050	1.852
	05/25/89	4	0.120	0.922	0.706	1.011
531	05/24/89	4	0.814	0.134	0.274	0.619
454	05/25/89	4	0.732	0.052	-0.090	1.259

3. Results: Button Presses

In the early SRI study[6], significant changes in alpha production were observed in response to an RS. The statistical evidence, however, did not indicate that the viewer was able to recognize an RS cognitively (i.e., the viewer's button presses relative to the RS did not exceed mean chance expectation).

In the current experiment, viewers 002, 009, and 372 were asked to press a button whenever they "perceived" an RS. The total number of stimuli during a session of 10 runs was not known in advance because of the randomization procedure. The null hypothesis is that the probability of a time interval having a stimulus is the same for those intervals with a button press as for those without a button press. In other words, the presence or absence of a stimulus is independent of the presence or absence of a button press. We tested this null hypothesis to determine if a viewer is cognitively aware of the RS.

In Table 2, the fractional hitting rate is $p_1 = A/(A+B)$, and the fractional missing rate is $p_2 = C/(C+D)$. The total number of 1-second intervals is $N = (A+B+C+D)$, and the total stimulus rate is $p_0 = (A+C)/N$.

Table 2

Data Schema for Interval Conditions

		Stimulus	
		Yes	No
Response	Yes	A	B
	No	C	D

Then, under the null hypothesis, the following statistic is approximately normally distributed with a mean of 0 and a variance of 1:

$$z = \frac{(p_1 - p_2)}{\sqrt{p_0(1-p_0)\left(\frac{1}{(A+B)} + \frac{1}{(C+D)}\right)}} .$$

Table 3 shows N, p_0, p_1, p_2, z, p-value, and the effect size, r, for the three sessions for which button-press data were collected. As in the earlier SRI study, there is no indication that the viewers were cognitively aware of the RS.

Table 3

Button Pressing Results

Viewer	N	p_0	p_1	p_2	z	p	r
002	1210	0.167	0.198	0.164	0.951	0.163	0.027
009	1280	0.091	0.068	0.094	−0.978	0.836	−0.027
372	1089	0.157	0.119	0.160	−0.996	0.840	−0.030

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

IV DISCUSSION AND CONCLUSIONS

We have found statistical evidence that the relative phase shift from –0.5 to 0.5 seconds of an RS are *not* characteristic of the data at large ($Z_s = 1.99$, $p \leq 0.024$, *effect size* $= 0.599$). The combined statistic for the PS indicates that the relative phase shift from –0.5 to 0.5 seconds of a PS are also *not* characteristic of the data at large ($Zs = 2.92$, $p \leq 0.002$, *effect size* $= 0.924$). Averaged across all viewers, the magnitude of the results, as indicated by their effect sizes of 0.599 and 0.924, respectively, is considered robust by accepted behavioral criteria defined by Cohen.[9]*

1. Root-Mean-Square Phase

Searching for a change of phase as a result of an RS is a natural extension of results quoted in the literature. For example, Rebert and Turner[5] report an example of photic driving (i.e., an extreme example of phase locking) at 16 Hz. In their work, a subject was exposed to a 16-Hz visual DS randomly balanced with no stimulus during 4-second epochs. The average power spectra showed approximately 10-Hz alpha activity during the no-light epochs, and a strong 16-Hz and no 10-Hz peak during the 16-Hz epochs.

One interpretation of their result is that the alpha rhythm was blocked, and the CNS "locked" on to the flashing stimulus. Eason, Oden, White and White,[10] report a phase-shift phenomenon when a rare stimulus, which is random relative to the internal alpha activity, is presented as a DS:

> "...*when a stimulus flash is presented, the resulting primary evoked response acts as a trigger stimulus which temporarily synchronized a certain percentage of the neural elements normally under the influence of an internal pacemaker. ... Desynchronization of the elements participating in the evoked response would occur as the elements are brought back under the influence of an internal pacemaker or are affected by neurons not involved in the response.*"

In other words, the internal alpha is momentarily interrupted by an external stimulus, and, in the absence of continuing external stimuli, returns back to its original frequency, but at a random phase relative to its pre-stimulus state.

To understand what would be expected in our experiment for the distribution of RMS phases during the Monte Carlo simulations, we examine a hypothetical case. Suppose that the viewer's alpha activity was a continuous wave at a single frequency. A phase change is computed between 500 ms before and 500 ms after each Monte Carlo "stimulus." Therefore, regardless of the entry point, the relative phase change would be zero, and the RMS phase over many such "stimuli" would also be zero.

Real alpha activity, however, is not continuous, Rather, it appears in bursts lasting from 100 to 5000 ms. Random Monte Carlo "stimuli" would sometimes occur within such bursts and sometimes near the edges. Thus, we would expect a nonzero RMS phase over many such "stimuli," but the individual relative phases would not be uniformly distributed. Depending upon the viewers' alpha characteristics, the distributions would be enhanced near zero RMS phase.

If we assume that Eason, et al., are correct, and that a phase shift is expected as a result of an RS, then the expected distribution of RMS phases is uniformly distributed on $[-\pi, \pi]$. In this case, the phase change is related to the relative timing between the external stimulus and the internal alpha—a completely random relationship. Thus, the variance of the RMS phases in the experimental condition should be larger than those computed during the Monte Carlo runs. Figure 13 is a schematic representation of these models.

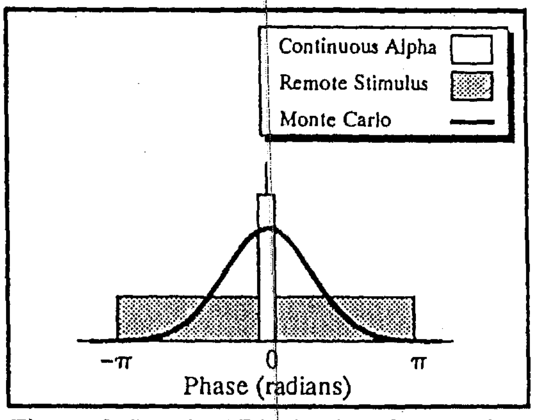


Figure 13 Idealized Distributions for Relative Phase Shifts

* Values of 0.1, 0.3, and 0.5 correspond to small, medium, and large effects, respectively.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

As a first step in testing these models, we computed the expected variance for the RMS phase, given that the individual phases are uniformly distributed on [π, π]. Using a Taylor Series expansion for RMS phase, the variance is given by:[11*]

$$\sigma_y^2 = \frac{\frac{3}{45}\pi^2}{n}\left[1 - \frac{1}{30n}\right] \quad (rad^2) \quad \text{, or}$$

$$\approx \frac{2160}{n} \quad (\deg^2),$$

where n is the number of individual phases.

Table 4 shows the viewer identification, the two-tailed z-score from Table 1, the number of RS, the theoretical variance for the RMS phase, the observed variance from the Monte Carlo runs of 500 passes each, and the X^2 and its associated p-value for a variance-ratio test.

Combining the X^2 across all 11 sessions gives an overall significant result ($X^2 = 5121.5, df = 5489, p \leq 0.0002$). This indicates that the Monte-Carlo-derived variances are significantly smaller than the theoretical variances based on uniformly distributed phases. The two viewers who demonstrated the largest z-scores (002 and 007) also show sharply reduced Monte Carlo variances.

Table 4

Comparison Between Monte Carlo Phases and Theory

I.D.	Z-Score (RS)	Number of RS	Variance of RMS Phase		X^2 df = 499	P-Value
			Theoretical	Observed		
009	-0.524	96	22.50	25.46	564.6	0.978
002	2.653	118	18.31	13.63	371.5	4.9×10^{-6}
	0.871	76	28.42	24.43	428.1	0.010
372	0.885	90	24.00	23.25	483.4	0.316
374	0.501	102	21.18	18.64	439.2	0.025
007	1.555	93	23.23	18.66	400.8	4.6×10^{-4}
389	-0.191	97	22.27	23.35	523.2	0.780
	-0.050	92	23.48	22.29	473/7	0.214
	0.706	98	22.04	20.22	457.8	0.093
531	0.274	101	21.39	21.05	491.1	0.408
454	-0.090	52	41.54	40.48	487.3	0.363

We must conclude that a uniform distribution for the phase is not a good assumption. To determine what the phase distribution was for the RS, we constructed histograms from the raw data.

Figure 14 shows the distribution of phases for the RS and Monte Carlo stimuli for viewer 002. While the RS distribution is enhanced near ±180 degrees and suppressed near 0 degrees compared to the Monte Carlo distribution, the differences are small ($X^2 = 10.62, df = 8, p \leq 0.224$) and, therefore, the random-phase model does not appear to be a good fit to the data for viewer 002 on his 25 September session.

Figure 15 shows the same distributions for viewer 007. In this case, the RS distribution is nearly uniform on [-180,180] degrees, but it differs only slightly from the Monte Carlo distribution ($X^2 = 9.47, df = 8, p \leq 0.304$).

From the data shown in Table 4, we see that the X^2 indicates significant overall differences between the theoretical and observed phase distributions. However, Figures 14 and 15 show that the differences between RS and Monte Carlo distributions are small. It is most probable, therefore, that the RS coupling to the CNS is weak, in general, and that the position of the sensor array is not necessarily optimized to sense the phase changes.

2. Viewer Dependencies

Viewers 002, 009, and 372 have produced consistent remote viewing results for many years—since 1972 for viewers 002 and 009, and since 1979 for viewer 372. Viewer 389 is a recent addition, and has produced examples of excellent remote viewing in the only experiment in which he has participated; however, he has produced significant results in another laboratory. Whereas viewer 002

* We thank Professor Jessica M. Utts, Statistics Department, University of California, Davis, California, for suggesting this approach.

produced the largest z-score (Z_s = *2.653*), viewer 009 produced the smallest (Z_s = *−0.524*). The combined effect size for the experienced viewers is 0.621, and is 0.559 for the inexperienced viewers. The difference is not significant.

There are two considerations that prevent drawing conclusions about the viewer dependence of the data. The number of independent samples is small, but the most compelling argument against drawing conclusions is that placement of the sensor array is a seriously confounding factor. As stated in Section II.2, we positioned the array in a location that maximized the response to a DS. This may not be the appropriate positioning for everyone. Indeed, it might not be optimal for anyone.

To determine if there were any "obvious" spatial dependencies that might indicate a more optimal array placement, we computed a complete set (all sensors) of Monte Carlo distributions for one session for viewer 002. Figure 16 shows the single-tailed p-values for the RMS phases for the RS and PS. They are displayed in the standard sensor-array configuration. The pattern for the RS suggests that a more optimal positioning of the array would be in the sensor 2-7 direction as indicated by an arrow in Figure 16.

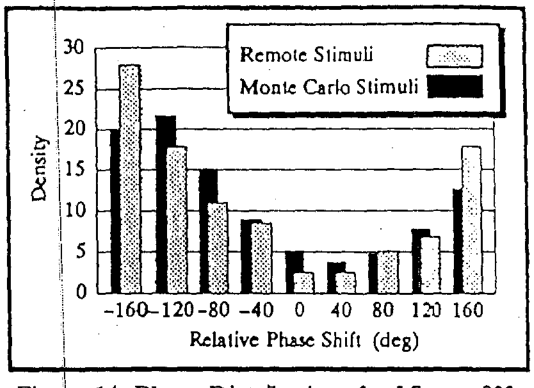


Figure 14 Phase Distributions for Viewer 002: 8/25/88

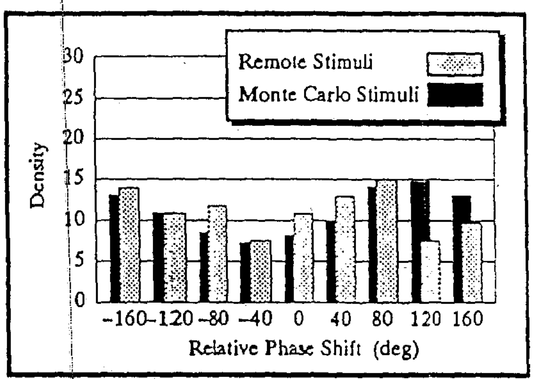


Figure 15 Phase Distributions for Viewer 007: 3/29/89

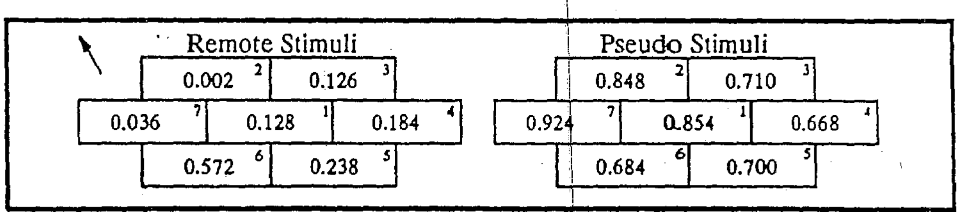


Figure 16 Phase p-values for Viewer 002: 8/25/88

3. Pseudo Stimuli

It was initially thought that the PS would act as a within-run control. The results indicate, however, that there was, on the average, a larger response to the PS than to the RS. While the difference was not significant, it is important to note that both of the responses are considered statistically robust (effect sizes of 0.599 and 0.924 for the RS and PS, respectively). A number of viewers' responses appear to produce phases on opposite sides of the Monte Carlo distributions (e.g., viewers 002 and 007), but there is no overall correlation between the RS and PS p-values.

A brief description of the hardware and software that is responsible for stimulus generation may help in understanding this outcome. The stimuli and their timing are imitated by an HP computer, but are controlled by an IBM PC. Each stimulus type has its own frame buffer within the PC. Our RS consists of a pattern of 1s and 0s that represent a sinusoidal grating in the center of an otherwise blank field. The PS pattern, a blank field that consists of all 0s, resides in a separate buffer. An interface board between the PC and a standard video monitor has its own internal frame buffer, which is automatically and continuously scanned at 30 Hz to provide a standard interleaved video signal. See Figure 17.

When the HP computer signals the PC to provide the appropriate stimulus, the following sequence of events are followed (see Figure 17):

(1) Phase locked to 60 Hz, the interface frame buffer is loaded with a copy of the appropriate stimulus frame buffer (either RS or PS).

(2) The interface board automatically sends this pattern interleaved at 30-Hz.

(3) After a preset time, approximately 100-ms in our experiment, the PC resets the interface frame buffer to zero (blank screen), and waits until another stimulus signal is received.

At the video monitor, the PS are indistinguishable from the between-stimuli blank screens. At the PC, however, the PS are distinguishable from the blank screen background, because the PC must copy a frame buffer (albeit all 0s) into the output frame buffer.

In our experiment, the RS and PS results were statistically identical, and independently, both were significantly different from the Monte Carlo distributions. This raises the question as to what constitutes the target stimulus. Our result is unexpected given the target was considered to be what was displayed on the remote monitor.

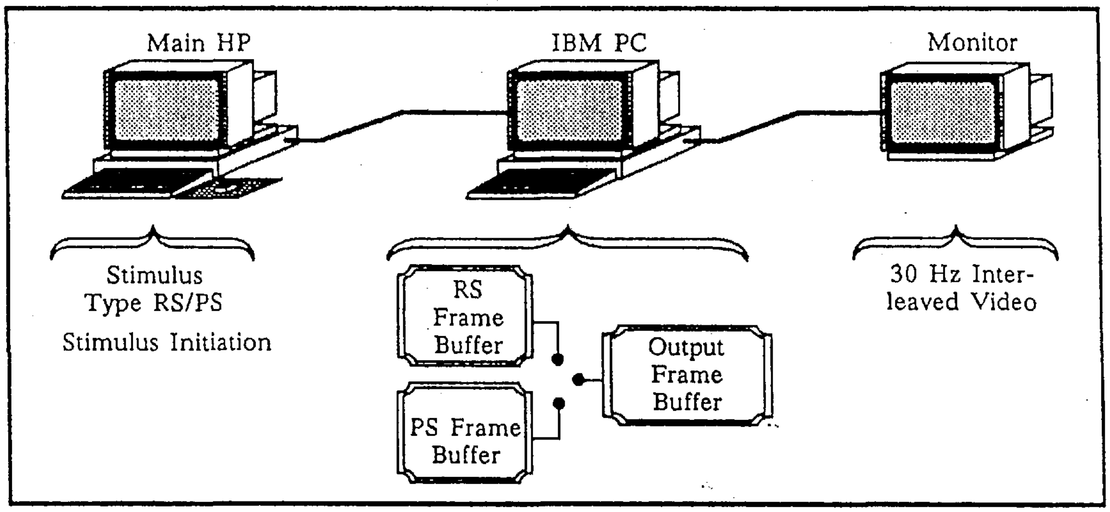


Figure 17 Sequence of Events for Stimuli Generation

It is conceivable that the internal activity of the PC, or its companion computer, was acting as an unintended target. If this were true, then there might be an electromagnetic (EM) coupling between the viewer's CNS and the internal electronic activity of the computers. It is well known that computers radiate EM energies at relatively high frequencies; for frequencies above 100 Hz, the shielded room is transparent. Analysis of the background runs (i.e., data collected in the absence of a sender or viewer) showed no EM coupling into the MEG electronics; therefore, it remains possible that the statistical effects we have seen are due to CNS responses to remote bursts of EM energy.

Let us *assume* that the overall RS and PS effects are meaningful. Since the PSs are *indistinguishable* at the monitor from the between-stimuli background but are *distinguishable* at the IBM PC, then the present experiment demonstrates that the source of stimuli is the IBM PC.

During the SRI/Langley Porter study in 1977, SRI developed an entirely battery operated stimulus generator as a special precaution against the possibility of system artifacts in the form of EM pickup. They reported significant CNS responses to remote stimuli, nonetheless.[6] Therefore, it remains possible that we have observed an anomalous information transfer.

Before further research is conducted, it is important to measure the EM radiation, and to see if it is of sufficient strength to be detected (by the appropriate hardware) in the shielded room.

By adjusting the PC program, the PS internal activity can be eliminated. It would be interesting to see if the similarity between the RS and PS results persists.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

REFERENCES

1. Dean, E. D., *International Journal of Neuropsychiatry*, Vol. 2, p 439, 1966.
2. Tart, C. T., *International Journal of Parapsychology*, Vol. 5, p 375, 1963.
3. Duane, T. D., and Behrendt, T., *Science*, Vol. 150, p. 367, 1965.
4. Cavanna, R., Ed., *Psi Favorable States of Consciousness*, Parapsychology Foundation, New York, 1970.
5. Rebert, C. S., and Turner, A., "EEG Spectrum Analysis Techniques Applied to the Problem of Psi Phenomena," *Physician's Drug Manual*, Vol. 6, Nos. 1-8, pp 82–88, 1974.
6. Targ, R., May, E. C., Puthoff, H. E., Galin, D., and Ornstein, R., "Sensing of Remote EM Sources (Physiological Correlates)," Final Report, Project 4540, SRI International, Menlo Park, CA, 1977.
7. Sutherling, W. W., Crandall, P. H., Cahan, L. D., and Barth, D. S., "The Magnetic Field of Epileptic Spikes Agrees with Intracranial Localizations in Complex Partial Epilepsy," *Neurology*, Vol. 38, No. 5, pp 778–786, May 1988.
8. Aine, C. J., George, J. S., Medvick, P. A., Oakley, M. T., and Flynn, E. R., "Source Localization of Components of the Visual-Evoked Neuromagnetic Response," Neuromagnetism Laboratory, Life Sciences and Physics Divisions, Los Alamos National Laboratory, Los Alamos, NM.
9. Cohen, J., *Statistical Power Analysis for the Behavioral Sciences* (rev. ed.), Academic Press, New York, 1977.
10. Eason, R. G., Oden, D., White, B. A., and White, C. T., "Visually Evoked Cortical Potentials and Reaction Time in Relation to Site of Retinal Stimulation," *Electroencephalography and Clinical Neurophysiology*, Vol. 22, pp 313–324, 1967.
11. Rice, J. A., *Mathematical Statistics and Data Analysis*, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, p 143, 1988.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Electrodermal Correlates of Remote Attention: Autonomic Reactions to an Unseen Gaze

William Braud, Donna Shafer, and Sperry Andrews

Mind Science Foundation

ABSTRACT

Anecdotal reports and a small number of prior laboratory studies suggest that persons are able to become aware of being stared at by another person, even in cases that preclude conventional sensory mediation. Previous studies assessed deliberate, conscious detection of staring by asking subjects to make verbal or motor "guesses" whenever they felt stared at by another, sensorily isolated, person. Such guesses were consistently accurate, but the effects were relatively weak. We hypothesized that stronger effects might be obtained if relatively "unconscious" autonomic nervous system activity were used as the indicator of staring detection, rather than conscious guessing. We reasoned that autonomic reactions might be less distorted by higher cognitive processes and might therefore provide a "purer" and more sensitive indicator. Two sets of sessions were conducted in which sympathetic nervous system activation was assessed by means of electrodermal monitoring during randomly interspersed remote staring and nonstaring (control) periods. The monitored subject was, of course, unaware of the number, timing, or pattern of these two types of periods. The possibility of sensory cueing was eliminated through the use of a closed-circuit television system for staring. Sixteen untrained subjects evidenced significant autonomic discrimination, becoming more activated during staring than during nonstaring periods. Sixteen subjects who had been extensively trained to become more aware of their interconnections with other people and less defensive about their connectedness also evidenced significant autonomic discrimination, but became more calm during staring than during nonstaring periods. Unconscious autonomic detection did indeed appear to yield stronger effects than did previous conscious verbal (or motoric) detection assessments.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

INTRODUCTION

Have you ever had the feeling that someone was staring at you from behind and, upon turning around, found you were correct? From time to time, most of us have had such a feeling, which appears to be a common part of the human experience. The feeling of "being stared at" is treated as a matter of course in countless novels by Tolstoy, Dostoyevsky, Hugo, Lawrence, Huxley, and many others. In surveys conducted in California (Coover, 1913) as early as 1913, 68-86 percent of respondents reported having had the experience on at least one occasion, and a more recent Australian survey (Williams, 1983) placed the figure at 74 percent. In a survey of San Antonio respondents recently completed as part of the present project, the figure was found to be approximately 94 percent.

Despite its widespread occurrence and familiarity, the staring experience has been subjected to surprisingly little scientific scrutiny. Is the presumed ability to detect an unseen gaze merely a superstition, a cultural myth without real substance, an over-inflation of coincidental occurrences, a response to subtle sensory cues, or could the experience have a paranormal component?

In 1898, Cornell University psychologist E. B. Titchener published a short article in Science in which he addressed the "feeling of being stared at." Titchener mentioned that he had conducted a series of laboratory experiments on this topic and that the experiments had yielded negative results; he reported no details whatsoever regarding those experiments. Titchener indicated that such experiments "have their justification in the breaking down of a superstition which has deep and widespread roots in the popular consciousness" (p. 897). He attempted to provide a psychological interpretation of the prevalence of the "staring" belief based on nervousness in social situations, attracting the attention of the starer, turning, and noticing the starer's gaze.

In 1912-1913, experimental research on staring detection was carried out by Stanford psychologist J. Edgar Coover. This was the first project that Coover undertook in his post as Fellow of Psychical Research--a position endowed (at $50,000 in 1911, then at $526,000 in 1918) by Thomas Welton Stanford, younger brother of the founder of the California school. In a paper published in the American Journal of Psychology (Coover, 1913), Coover reported the results of a study in which each of 10 subjects made 100 guesses of whether he or she was being stared at by an experimenter seated behind the subject in the same room; the subject kept his or her eyes closed and "shaded with one hand." The staring versus nonstaring schedule was determined by tossing a die. The duration of a staring or nonstaring trial was 15-20 seconds; the 100 trials were distributed over 3 to 4 hourly sessions that were spaced one week apart. Overall, the subjects' accuracy of guessing did not depart significantly from chance. Coover discussed qualitative differences in the subjects' imagery and subjective impressions that he thought were correlated with the degree of confidence or certainty of their guesses, but did not substantiate his conclusions with quantitative data. He interpreted his findings as support for Titchener's claim that the belief in staring detection was empirically groundless.

In 1959, J. J. Poortman of Leyden University (Netherlands) reported a preliminary staring detection study in the Journal of the (British) Society for Psychical Research. Poortman himself served as a subject for 89 trials

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

(distributed over a 13 month period) in which he attempted to guess whether or not he was being stared at by another experimenter; the same person served as experimenter throughout the tests. Poortman was seated in a separate room adjoining that of the starer, with his back to the starer. Staring and nonstaring trials were of 2-5 minutes duration and were randomly scheduled by means of card shuffles. Portman achieved a 59.55 percent accuracy rate which he called "suggestive and highly promising" although not significant. (A re-analysis of Poortman's data by the present authors yields a one-tailed p = .04). Poortman also provided several interesting observations of psychological conditions that appeared, in his own experience, to facilitate or to impede accurate staring detection.

In the Coover and Poortman experiments, test conditions were poorly controlled. The subject and the starer were in the same room or in open adjoining rooms and the subject could have discriminated staring from nonstaring periods by means of subtle, unintentional auditory cues. This cueing possibility was eliminated in two recent studies by Peterson (1978) and Williams (1983).

Donald Peterson conducted two preliminary pilot studies and a third formal experiment as part of a final year project for his M.A. degree in the Psychology Department of the University of Edinburgh; these experiments are still unpublished. The pilot studies were relatively informal and were conducted in order to ascertain effective procedures that would later be used in the formal experiment. That formal experiment involved nine starer-staree pairs of participants. The starer and staree occupied separate, adjacent, closed cubicles. Special lighting and the use of one-way mirrors permitted visual access in one direction only--i.e., the starer could see but could not be seen by the staree. Isolation was increased further by requiring the staree to listen to sound-masking white noise through headphones. The staree pressed a pushbutton whenever he or she felt "stared at"; the button presses marked a chart recorder and provided "time on target" measures. The staring and nonstaring periods were scheduled randomly by means of special equipment. The actual test trials were preceded by brief training periods in which the staree received feedback in an attempt to develop an appreciation of internal cues that might be associated with staring detection. The members of each dyad reversed roles during the experiment so that each person had an opportunity to serve as both starer and staree. There were 36 experimental sessions overall, each of six minutes duration; each session contained three 30-second staring periods. Analysis of results indicated significantly accurate detection of staring (p = .012, two-tailed).

The experimental design was improved even further by Linda Williams, who conducted her studies in the Psychology Department of the University of Adelaide (South Australia) as part of her B.A. honors thesis work. Williams presented portions of this still unpublished work at a 1983 conference. Williams (1983) provided excellent sensory isolation of her starers and starees by stationing them in separate, closed rooms 60 feet apart. Rather than using a one-way mirror, the starer watched the subject by means of a closed-circuit video camera/monitor arrangement. Twenty-eight starees participated in the study and indicated their staring detection guesses by means of button presses. Each staree experienced 52 12-second staring trials and 52 12-second nonstaring trials; the two trial types were scheduled randomly by means of signalling tapes created on the basis of random numbers. Conventional measures of accuracy, as well as sensitivity measures (d') derived from signal detection

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

theory, yielded significant results (p = .04, one-tailed).

Table 1

Summary of Previous Staring Detection Experiments

Investigator	Affiliation	Design Features	Scoring Rate	Effect Size
E.B. Titchener (1898)	Cornell University	no data reported	---	---
J.E. Coover (1913)	Stanford University	same room	50.20%	.004
J.J. Poortman (1959)	Leyden University	adjoining rooms	59.55%	.18
D.M. Peterson (1978)	Edinburgh University	one-way mirrors	54.86%	.42
L. Williams (1983)	Adelaide University	closed-circuit TV	51.31%	.32

Three of the four empirical studies reviewed above yielded suggestive evidence that persons are able to consciously discriminate periods of staring from those of nonstaring, even in the two cases in which the possibility of subtle sensory cues was eliminated. In fact, an examination of the tabulated results (see Table 1) reveals that scoring actually improved as test conditions were made more stringent, especially if success is measured in terms of effect size (defined as z-score or equivalent z-score divided by the square root of the number of contributing score units; see Rosenthal, 1984).

The effect, however, although consistent, was not particularly striking. A plausible reason for this is that the testing method used in these studies was not the most appropriate one. The laboratory experiments were designed to encourage deliberate conscious guessing in order to identify staring periods. Such a procedure would be expected to maximize possible cognitive interferences and distortions of subtle internal staring-related cues; it would be difficult or impossible for the staree to avoid the use of guessing strategies, response biases, intellectual analysis and interpretation, etc. In the everyday life context, on the other hand, staring detection frequently takes the form of spontaneous unconscious behavioral and bodily changes. Often, such changes are reported to be rich in physiological content (e.g., tingling of the skin, prickling of neck hairs) and automatic movements (e.g., spontaneous head-turning, unplanned glances). Higher cognitive functions seem to play minor roles in these staring detection contexts.

On the basis of these considerations, it was hypothesized that an experimental design based upon more unconscious autonomic nervous system reactions might be more sensitive to staring detection than would one based upon conscious motor

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

guessing. Therefore, we designed an experiment in which we would be able to monitor sympathetic autonomic nervous system activity (using electrodermal recording techniques) in the staree during staring and nonstaring periods to determine whether those periods could be unconsciously differentiated. As in the Williams study, separate closed rooms and a closed circuit television system would be used to eliminate conventional communication channels between starer and staree. We also sought to compare the autonomic staring detection ability of two groups of subjects. One group (tested in Phase 1) consisted of untrained subjects. Another group (tested in Phase 2) consisted of subjects who had undergone special training designed to increase their understanding and experience of human interconnectedness and to help them deal with their psychological resistances to feeling interconnected with others.

METHOD

Subjects

Thirty-two subjects participated as starees in this experiment. The subjects were unselected (perhaps a better description would be "self-selected") persons from the local community who were interested in our experiments and who had become aware of the Mind Science Foundation's work through media information (local radio, newspaper, and newsletter descriptions) and information from previous subjects. The participants ranged in age from 22 to 71 years; there were 24 females and 8 males. The 32 subjects were tested in two phases. Phase 1 involved 16 subjects who were "untrained." Phase 2 involved 16 subjects who had been ("self-") selected from the same general subject pool; these subjects, however, were tested following their participation in a "connectedness training" program (see below) conducted by the third author (S.A.). The second author (D.S.) played the role of both experimenter and starer throughout the experiment; she, too, participated in the connectedness training following Phase 1 but before Phase 2.

Apparatus

The experimental apparatus consisted of silver/silver chloride palmar electrodes, a skin-resistance amplifier, and an analog-to-digital converter interfaced with a microcomputer. This equipment was identical to that described in previous reports (Braud & Schlitz, 1983; Schlitz & Braud, 1985). A closed circuit television system was added, so that the staree could be viewed by a distant starer without sensory cueing. A color video camera (Hitachi camcorder VM-2250) was set up on a tripod in the staree's room. The camera's radio frequency (RF) output was boosted by a 10 dB amplifier then conveyed via heavy duty 300-ohm impedance twin-lead cable to a 19-inch color TV monitor (Sony Trinitron KV-1114) situated in the starer's room.

Procedure

The experimenter met with the subject in the Foundation's library, where the subject completed the Myers-Briggs Type Indicator (Form F: see Briggs & Myers, 1957), a 55-item personal history survey (Participant Information Form [PIF]) developed at the Psychophysical Research Laboratories (Psychophysical Research Laboratories, 1983), and a brief staring questionnaire. The staring questionnaire asked whether the subject had ever felt an unseen person staring at him or her and whether such an experience took the form of a physical sensation or a conscious thought; the questionnaire also asked that the subject

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

describe the experience. After completing these assessments, the subject was taken to the starer's room, shown the television monitor, and informed of the details of the procedure.

Next, the experimenter led the subject to the staree's room, which was located in an entirely different suite area across an outside corridor. The two rooms were separated from each other by two inner hallways, an outer corridor, and four closed doors. Conventional sensorimotor communication between these two rooms, under the conditions of the experiment, was not possible. The staree's room was brightly illuminated by means of overhead fluorescent lights. The camera was mounted on a tripod six feet away from the subject's chair, at eye level, and at an angle of approximately 45 degrees left of center (from the subject's point of view). The camera's zoom lens was set so that the subject's shoulders, neck, and head would be visible on the monitor in the starer's room.

The subject was seated in a comfortable recliner chair (which remained in an upright position throughout the experiment), and the experimenter attached two silver/silver chloride electrodes (7 mm diameter) filled with partially conductive electrode gel to the subject's left palm by means of adhesive electrode collars. The subject was asked to sit quietly for the next 20 minutes, to refrain from unnecessary movements (especially of the left hand and arm), and to think about whatever he or she wished during the experiment. The subject was told that the camera would be on throughout the 20-minute session, but that the experimenter would watch the monitor only at certain randomly determined times during the experiment; at those times, the experimenter would stare intently at the subject's image on the distant monitor and would attempt to gain the subject's attention. The subject was asked to "keep in the back of your mind a gentle wish that the experiment will be successful." The experimenter then left the subject alone in the staree room and went to the distant starer's room, closing all doors behind her.

In the starer's room, the experimenter recorded the subject's basal skin resistance, then retrieved from a hidden location a sealed, opaque envelope that contained the random staring/nonstaring periods sequence for that session. Thirty-two such envelopes had been prepared previously by W.B., who had used the random algorithm of G W - BASIC to generate the random sequence of the 10 staring and 10 nonstaring periods for each session. In a hidden location known only to him, W.B. kept his own copies of the 32 random sequences. The experimenter started the microcomputer program that controlled the timing of the various events of the experiment and recorded the subject's electrodermal activity during each of 20 30-second recording periods. Each of the 20 recording periods was signalled by a low-pitched tone (audible only to the experimenter, through headphones); a 30-second rest period followed each recording period. The experimenter consulted the contents of the session envelope to learn which of the 20 recording periods were to be devoted to staring and which were to serve as the nonstaring, control periods. If the random sequence indicated a staring period, the experimenter swiveled her chair around so that it faced the TV monitor; she stared intently at the subject's monitor image throughout the 30-second recording period. During nonstaring periods, the experimenter kept her chair turned away from the monitor, so that she could not see the monitor's screen; she busied her mind with things unrelated to the experiment. All reflective surfaces had been carefully covered so that inadvertent glimpses of the monitor screen were not possible. The experimenter did not receive electrodermal feedback of any type during any phase of the experiment.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

At the end of the experimental session, the computer printed the electrodermal results for each of the 20 recording periods. These scores represented the integrated then averaged values of all spontaneous skin resistance reactions that occured during each recording period (see Braud & Schlitz, 1983, for details). The experimenter carefully tore off the paper printout, folded it without looking at the electrodermal scores, and filed it away in a special location. Then, she went to the staree's room and discussed the experiment with the subject in general terms. Neither experimenter nor subject had any knowledge of the numerical results for the session. Only after all 32 sessions had been completed did W.B. analyze the results and give the experimenter feedback. The experimenter later provided feedback to those subjects who requested it.

The first 16 subjects who participated in Phase 1 of the study had no special preparation. The 16 starees of Phase 2, however, had received "preparation" in the form of participation in "connectedness training" workshops conducted at Mind Science Foundation by the third author (S.A.). Workshop participants had spent approximately 20 hours engaging in intellectual and experiential exercises involving feelings of interconnectedness with other people and dealing with their own psychological resistances to merging with others. The workshop began with a group viewing of a videotape based on Peter Russell's book, The Global Brain (Russell, 1983). This was followed by discussions of the videotape, lectures by S.A., and experiential exercises in which participants became increasingly comfortable and adept at "connecting" with each other. The latter took the form of staring into another person's eyes for long periods of time, becoming comfortable with this, observing how one's physiological reactions came to more closely resemble those of the other, and conversing and retrieving information while maintaining eye contact (rather than averting the gaze upward or sideward, as would usually occur during memory retrieval and cognitive processing; see Bakan, 1980). Individual and group discussions were devoted to learning about and dealing with psychological resistances that interfered with the process of connectedness or with feelings of merging with another person. The experimenter/starer for the present study (D.S.) actively participated in all workshop sessions. All workshop participants were aware that the workshop would be followed by an experiment involving physiological detection of staring, but were not aware of any more details of the study than were the 16 untrained subjects of Phase 1.

In the present study, we simply explored the possible effects of connectedness training. We did not make any predictions about scoring direction in the two phases, and therefore planned to use two-tailed tests in their evaluations. It could be argued that the training may have increased the sensitivity of Phase 2 subjects to whatever effects occurred in Phase 1. Alternatively, the training may have resulted in a qualitatively different reaction pattern in Phase 2.

RESULTS

For each of the 32 sessions, a total score was calculated for all 20 recording periods (10 staring and 10 nonstaring). This total score was divided into the sum of the electrodermal activity scores for the 10 staring (S) periods; the process was repeated for the 10 nonstaring (N) periods. In the absence of a remote staring effect, these two ratios [S/(S + N), N/(S + N) should approximate 50 percent. A remote staring effect would be indicated by a significant departure of the scores from the 50 percent mean chance expectation. Single-mean tests were used to assess the departure of the ratios

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

from M.C.E. (50 percent). This is equivalent to calculating dependent (matched) t tests to compare the raw scores for each subject for staring versus nonstaring periods. We have consistently used ratio scores in our various projects as a method of "standardizing" scoring so that scoring magnitude could be more meaningfully compared for the different response systems that we measure.

First, an analysis was performed on the staring/total activity ratios of the 16 untrained subjects of Phase 1. A single mean t test indicated that the 16 untrained starees exhibited significantly greater spontaneous electrodermal activity during staring periods than during nonstaring, control periods. The mean percent electrodermal activity for staring periods was 59.38%, rather than the 50.00% expected by chance. The single mean t test comparing the 16 staring period percent scores with 50% M.C.E. was 2.66 which, with 15 degrees of freedom, has an associated two-tailed $p = .018$, and a calculated effect size = .59; the 95% confidence interval is bounded by the values 51.86% and 66.90%. Thus, these subjects were significantly activated (in terms of sympathetic autonomic activity) by remote staring, compared to the nonstaring, control periods.

Next, a parallel analysis was performed on the scores for the 16 Phase 2 subjects who had experienced connectedness training prior to their experimental sessions. A single mean t test indicated that the 16 trained starees exhibited significantly less spontaneous electrodermal activity during staring periods than during nonstaring, control periods. The mean percent electrodermal activity for staring periods was 45.45%, rather than the 50.00% expected by chance. The single mean t test comparing the 16 staring period percent scores with 50% M.C.E. was 2.15 which, with 15 degrees of freedom, has an associated two-tailed $p = .048$, and a calculated effect size of .50; the 95% confidence interval is bounded by the values 40.94% and 49.95%. Thus, these trained subjects were significantly calmed (in terms of sympathetic autonomic activity) by remote staring, compared to the nonstaring, control periods.

If the scores for the subjects of the two Phases are directly compared by means of an independent groups t test, a significant difference is found between the untrained and the trained subjects ($t = 3.39$, $df = 30$, $p = .002$, two-tailed).

Secondary analyses were performed to test the equivalence of the Phase 1 and Phase 2 subjects in terms of their personality (MBTI) and physiological (electrodermal activity) characteristics; a summary of these analyses is presented in Table 2. For the MBTI scores, group means are presented for the continuous scores of the extraversion/introversion, sensing/intuition, thinking/feeling, and judging/perceiving dimensions. A score of 100 represents the midpoint of each continuum. Scores less than 100 indicate tendencies toward extraversion, sensing, thinking and judging; scores greater than 100 indicate tendencies toward introversion, intuition, feeling and perceiving. For electrodermal activity scores, group means are given for the sum of spontaneous skin resistance responses integrated over all 20 recording epochs (total SRR) and for the subjects' initial basal skin resistance (BSR) in ohms. High total SRR scores and low BSR scores are associated with increased sympathetic autonomic arousal, whereas low total SSR scores and high BSR scores are associated with decreased sympathetic arousal. Analyses indicated that the Phase 1 and Phase 2 groups did not differ significantly on any of these six measures.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 2

Group Means and Statistical Comparisons of Personality and Physiological Characteristics of Phase 1 and Phase 2 Subjects

	MBTI Continuous Scores				Electrodermal Activity	
	E I	S N	T F	J P	Total SRR	BSR
Phase 1 (Untrained)	87.12	123.00	106.75	117.00	605.19	343,506
Phase 2 (Trained)	98.75	133.69	98.06	103.50	656.06	289,047
t	1.49	1.44	1.23	1.52	0.54	0.82
p (2t)	.15	.16	.28	.14	.60	.42

We are now able to supplement the findings previously summarized in Table 1 with the results of the present investigation (see Table 3).

Table 3

Summary of Present Autonomic Staring Detection Experiments

Investigator	Affiliation	Design Features	Scoring Rate	Effect Size
Present study (1990)	Mind Science Foundation	closed circuit TV autonomic measures		
untrained subjects			59.38%	.59
trained subjects			54.55%	.50

A more detailed statistical summary of all relevant staring detection research is presented in Table 4. If effect size is taken to be the most appropriate measure of the strength of an obtained outcome, it appears that the autonomic recording method of the present study does indeed yield stronger results than do the conscious guessing measures of staring detection used in previous studies.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Table 4

Statistical Summary of All Staring Detection Experiments

Study	test	p (1-t)	z	n	Effect Size
Coover (1913)	z = 0.126	.4499	0.126	1000	.004
Poortman (1959)	z = 1.70	.044	1.70	89	.18
Peterson (1978)	t = 2.648	.006	2.51	36	.42
Williams (1983)	t = 1.77	.044	1.70	28	.32
Present: untrained subjects	t = 2.66	.009	2.37	16	.59
Present: trained subjects	t =-2.15	.976	-1.98	16	-.50

The statistical values of Table 4 may be used in a preliminary meta-analysis of all staring detection studies reported to date. The table lists the statistical test presented in the original report, the one-tailed p value associated with that test, the z-score equivalent of that one-tailed p value, the number of units contributing to the analysis, and the effect size (calculated by dividing the equivalent z score by the square root of n). Combining all six tabulated entries yields a mean z = 1.07, a Stouffer z = 2.62 (with associated one-tailed p = .0044), and a mean effect size = .17. The Stouffer z procedure, a preferred method for combining probabilities of several studies testing essentially the same hypothesis, is described by Rosenthal (1984); this source also provides an excellent discussion of various effect size measures.

A comment is necessary regarding the last entry of Table 4. In the present research, we sought to determine whether the subjects would autonomically discriminate the staring from the nonstaring (control) periods; indeed, they were able to do this in both Phases. No prediction was made regarding the direction of their differential autonomic response, i.e., whether their electrodermal activity would be greater or less during staring periods (compared with nonstaring periods). For this reason, two-tailed tests were used for each Phase, and results for both Phases were "successful." However, for purposes of meta-analysis, it is customary to use only one-tailed tests and p-values in the tabulations. It is also customary to use a negative sign for a result that is inconsistent with the bulk of the results (see Rosenthal, 1984, p. 95). We have followed this convention when entering the results for the trained subjects (Phase 2) of this study. This provides a conservative estimate of overall results, because the autonomic discrimination of the trained, Phase 2 subjects was just as effective as that of the untrained, Phase

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

1 subjects, but happened to be in the "calm" as opposed to the "active" direction. This "reversal" of direction becomes understandable when considered in relation to the nature of the training experienced by these subjects (see below).

It should also be pointed out that the units of analysis for the effect sizes reported in Table 4 are not comparable for all studies. Effect sizes for the first two studies (those of Coover and of Poortman) are based upon trial units, whereas those of the remaining studies are based upon subject units. These differences should be borne in mind when evaluating these effect sizes.

DISCUSSION

Prior research had yielded suggestive evidence that persons were able to discriminate staring and nonstaring periods by means of deliberate, conscious guesses. The aim of the present project was to determine whether staring and nonstaring periods could be differentiated by means of more "unconscious" physiological reactions. The electrodermal activity differences between staring and nonstaring periods indicated that such differentiation could indeed occur. We chose to measure spontaneous electrodermal fluctuations (i.e., changes in skin resistance reactions) because such measurements are easy to make, are sensitive indicators, and are known to be useful peripheral measures of the activity of the sympathetic branch of the autonomic nervous system. The occurrence of many or of high amplitude skin resistance reactions (SRRs) is symptomatic of increased sympathetic activation or arousal, which may in turn reflect increased emotionality (see Braud, 1981; Edelberg, 1972; Prokasy & Raskin, 1973). On the other hand, the occurrence of few or of low amplitude SRRs indicates decreased sympathetic activation or arousal, which may in turn reflect decreased emotionality and, therefore, a greater degree of emotional and mental quietude or calmness.

The results of both phases of this study indicate reliable autonomic discrimination of staring and nonstaring periods, and the relatively large effect sizes suggest that autonomic detection may be a more powerful method than conscious guessing for the detection of staring effects. Phase 1 findings suggest that the starees were more activated during the staring than during the nonstaring epochs. Phase 2 findings suggest that those starees were more calm during staring than during nonstaring periods. This latter finding does in fact make sense in view of the training experienced by the Phase 2 subjects. That training was designed to allow persons to become more comfortable with staring and with "connecting" with other people, and permitted the trainees to reduce at least some of their defenses or resistances to staring, being stared at, and "merging" with another person. In the course of their training, many trainees reported that their staring encounters became quite positive and pleasant interactions, and they expressed disappointment when the encounters ended. We speculate that similar processes may have occurred during remote staring: The trained subjects of Phase 2 may have "missed" contact with and the attention of the starer (with whom they had become increasingly familiar during the course of the training), and may have become more relaxed and calm when that attention was provided, albeit in remote form, during the staring epochs of the experiment. A useful analogy for the reader might be the alarm and distress that occur upon the removal of an imprinted object from the environment of an imprinted precocial fowl or other organism, and the distress-reduction that occurs when the imprinted object is re-introduced (see Bateson, 1966; Johnson, 1989; Peoch, 1988; Ratner & Denny, 1964). For the Phase 1

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

starees, who did not have the benefit of the connectedness training, being stared at (even in its remote form) may have been experienced in a more typical way, i.e., as threatening (see Argyle, 1975) and sympathetically activating (rather than calming).

The above comments apply not only to the starees, but also to the starer. Although she attempted to behave identically and maintain identical attitudes in the two phases, when she began Phase 2, the starer (D.S.) had, of course, participated actively in the connectedness training and may well have been more comfortable and relaxed about merging with her subjects in Phase 2 than she had been in Phase 1. This increased comfort and relaxation could have been reflected in the "calming" direction of the Phase 2 results.

It is possible, of course, that the different patterns of findings for Phase 1 and Phase 2 are manifestations of a differential effect contributed by the experimenter/starer's experience of two juxtaposed sets of experimental sessions (see Rao, 1965, 1985). Alternatively, it is possible that the different patterns are contributed by a nonequivalence of characteristics of the subjects in the two phases. This latter hypothesis is not a convincing one, since the subjects for the two phases came from the same general participant pool and did not appear to differ importantly in terms of PIF characteristics, MBTI profiles, or overall electrodermal reactions (either basal skin resistance or total electrodermal activity for the session, both of which reflect general arousal level, nervousness, etc.). It would appear that the participants in the two phases were sufficiently similar in their initial characteristics to rule out differences attributable to those factors alone. The participants' training (that of the starees and of the starer in Phase 2) appears to have been more critical in determining the results. Further research will clarify these issues.

In addition to its relevance to staring detection, the present findings contribute to our understanding of processes involved in our ongoing series of "biological psychokinesis" (bio-PK) or "remote mental influence" experiments. In those experiments, we have found that "influencers" can produce directional changes in electrodermal activity and other reactions in spatially distant living target systems through the use of intentional and imagery strategies (see Braud & Schlitz, 1989, and Braud, Schlitz & Schmidt, 1989, for summaries of these findings). The findings of the present staring-detection study indicate that the mere focusing of attention or awareness upon a remote living system may perturb that system in a nontrivial manner, even in the absence of deliberate attempts to influence a specific aspect of the system in a particular direction. It is significant that such attention-mediated effects can occur (as they did in the present research) without the provision of real time sensory feedback to the attention source. Our bio-PK experiments, therefore, would appear to include two important components: (a) a specific, directional influence (determined by the experiment's protocol), as well as (b) a more general, nonspecific effect contributed by the deployment of the influencer's pure attention or awareness upon the target system. Similar effects of attention per se have been reported in other experiments with living systems (e.g., Pleass & Dey, 1985). The effects of components (a) and (b) would, of course, be expected to be modulated by the predispositions of the systems or persons being influenced in these experiments, and the "influencees" could provide additional components that could combine additively or multiplicatively with influences (a) and (b).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Autonomic Staring Detection

It would be of great interest to repeat the present experiments using instrumentation that permitted real-time monitoring of brain electrical or magnetic activity, instead of or in addition to the peripheral autonomic measures used here. Such a project, involving five independent, cooperating brain research laboratories, has been designed by S.A. and is currently being pursued as a collaborative research effort. Spatio-temporal topographical maps of electroencephalographic or magnetoencephalographic activity, recorded simultaneously in starer and staree, could reveal subtle and complex interactions that have heretofore evaded detection. Several suggestive investigations of this type have already been reported (Grinberg-Zylberbaum & Ramos, 1987; Grinberg-Zylberbaum, Delafor & Sanchez-Arellano, 1990; Orme-Johnson, Dillbeck, Wallace & Landrith, 1982). The newly developed brain electrical activity mapping (BEAM) technology, which employs traditional EEG scalp electrodes and amplification, but intensive real time computer processing (fast Fourier transformations and extensive high resolution color topographical displays), promises to be the instrumentation of choice for such future investigations. BEAM technology allows the display of brain maps that rival those produced by computerized axial tomography (CAT scans) or positron emission tomography (PET scans), but without the discomfort, invasiveness, potential risks, or expense of the latter two methods.

The present finding of a correlation between the attentional processes of one person and the physiological activity of another, spatially separated, person (along with the findings of numerous "bio-PK" investigations and, indeed, of the entire body of evidence from parapsychology and psychical research) is consistent with a conceptualization in which people are viewed as sharing a special interconnectedness. Such an interpretation is also consistent with a view of mind as unitary and nonlocal (see Dossey, 1989, for the most recent, and perhaps the single most satisfactory, exposition of this theme).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

REFERENCES

Argyle, M. (1975). Bodily Communication. New York: International Universities Press. Pp. 229-250.

Bakan, P. (1980). Imagery, raw and cooked: A hemispheric recipe. In J. Shorr, G. Sobel, P. Robin & J. Connella (Eds.), Imagery: Its Many Dimensions and Applications. New York: Plenum. Pp. 35-53.

Bateson, P. P. G. (1966). The characteristics and context of imprinting. Biological Review, 41, 177-220.

Braud, W. G. (1981). Psi performance and autonomic nervous system activity. Journal of the American Society for Psychical Research, 75, 1-35.

Braud, W. G. & Schlitz, M. J. (1983). Psychokinetic influence on electrodermal activity. Journal of Parapsychology, 47, 95-119.

Braud, W. G. & Schlitz, M. J. (1989). A methodology for the objective study of transpersonal imagery. Journal of Scientific Exploration, 3, 43-63.

Braud, W. G., Schlitz, M. J. & Schmidt, H. (1989). Remote mental influence of animate and inanimate target systems: A method of comparison and preliminary findings. Proceedings of Presented Papers: The 32nd Annual Parapsychological Association Convention, San Diego, CA, pp. 12-25.

Briggs, K. C. & Myers, I. B. (1957). Myers-Briggs Type Indicator Form F. Palo Alto, CA: Consulting Psychologists Press.

Coover, J. E. (1913). The feeling of being stared at. American Journal of Psychology, 24, 57-575.

Dossey, L. (1989). Recovering the Soul: A Scientific and Spiritual Search. New York: Bantam.

Edelberg, R. (1972). Electrical activity of the skin: Its measurement and uses in psychophysiology. In N. Greenfield & R. Sternbach (Eds.), Handbook of Psychophysiology. New York: Holt, Rinehart & Winston. Pp. 367-418.

Grinberg-Zylberbaum, J., Dalafor, M. & Sanchez-Arellano, M. E. (1990). The "transferred" potential in the human brain. International Journal of Neuroscience, in press.

Grinberg-Zylberbaum, J. & Ramos, J. (1987). Patterns of interhemispheric correlation during human communication. International Journal of Neuroscience, 36, 41-55.

Johnson, M. H. (1989). Imprinting and anpsi: An attempt to replicate Peoch, 1988. Journal of the Society for Psychical Research, 55, 417-419.

Orme-Johnson, D., Dillbeck, M. C., Wallace, R. K. & Lindrith, G. S. (1982). Intersubject EEG coherence: Is consciousness a field? International Journal of Neuroscience, 16, 203-209.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Peoch, R. (1988). Chicken imprinting and the tychoscope: Anpsi experiment. Journal of the Society for Psychical Research, 55, 1-9.

Peterson, D. M. (1978). Through the looking-glass: An investigation of the faculty of extra-sensory detection of being stared at. Unpublished M.A. thesis, University of Edinburgh, Scotland.

Pleass, C. M. & Dey, N. D. (1985). Using the Doppler effect to study behavioral responses of motile algae to psi stimulus. Proceedings of Presented Papers, Volume 1: The 28th Annual Parapsychological Association Convention, Tufts University, Medford, MA, pp. 373-405.

Poortman, J. J. (1959). The feeling of being stared at. Journal of the Society for Psychical Research, 40, 4-12.

Prokasy, W. F. & Raskin, D. C. (1973). Electrodermal Activity in Psychological Research. New York: Academic Press.

Psychophysical Research Laboratories (1983). 1983 PRL Annual Report. Princeton, NJ: Psychophysical Research Laboratories.

Rao, K. R. (1965). The bidirectionality of psi. Journal of Parapsychology, 29, 230-250.

Rao, K. R. (1985). Meta-analysis of the "Differential Effect" in ESP Testing. Unpublished manuscript, Institute for Parapsychology, Durham, NC.

Ratner, S. C. & Denny, M. R. (1964). Comparative Psychology: Research in Animal Behavior. Homewood, IL: Dorsey Press. Pp. 351-368.

Rosenthal, R. (1984). Meta-analytic Procedures for Social Research. Beverly Hills, CA: Sage Publications.

Russell, P. (1983). The Global Brain. Los Angeles, CA: J. P. Tarcher.

Schlitz, M. J. & Braud, W. G. (1985). Reiki-Plus natural healing: An ethnographic/experimental study. Psi Research, 4, 100-123.

Titchener, E. B. (1898). The feeling of being stared at. Science, 8, 895-897.

Williams, L. (1983). Minimal cue perception of the regard of others: The feeling of being stared at. Paper presented at the 10th Annual Conference of the Southeastern Regional Parapsychological Association, West Georgia College, Carrollton, GA, February 11-12. [See also Williams, L. The feeling of being stared at: A parapsychological investigation. Unpublished manuscript, n.d.]

Mind Science Foundation
8301 Broadway, Suite 100
San Antonio, Texas 78209

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

REMOTE VIEWING: FROM WHAT TIME FRAME DOES THE INFORMATION ORIGINATE?

by

Nevin D. Lantz, Edwin C. May, and Thomas Piantanida
SRI International, Menlo Park, California

ABSTRACT

In order to determine from what time frame remote viewing information originates, we have examined the role of precognition and feedback on remote viewing quality. We displayed the feedback tachistoscopically at the end of an otherwise standard remote viewing protocol. The cognitive awareness of the feedback experience was minimal, and 2 of the 8 intensities used for visual display of the feedback were below subliminal threshold. We hypothesized a number of possible relationships between feedback intensity and remote viewing quality, including one based on precognition (i.e., the remote viewing data originated from the future feedback). Four viewers contributed 40 trials each (5 at 8 different intensity bands). Using a sum-of-ranks statistic, two viewers produced independently significant evidence of remote viewing ($p \leq 3.5 \times 10^{-6}$, effect size = 0.711, and $p \leq 0.012$, effect size = 0.357, respectively). None of the data showed significant correlation of feedback intensity with remote viewing quality. This result is discussed with regard to precognition in general and the troublesome unfalsifiablity aspect of truly goal-oriented precognition.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

I INTRODUCTION

One model of remote viewing, which is based on precognition, is that the data originate from the future feedback to the viewer. If we are to understand the process of remote viewing, we must examine and understand the process of precognition. Unfortunately, precognition itself is not understood. For some experimenters, it is simply a "clean" methodology for conducting free-response experiments, while others have attempted to understand its mechanisms. Our interest was to explore the properties of precognition by examining the relationship between feedback and remote viewing quality.

For the purposes of this paper we have assumed that precognition is real. While there are many examples in the parapsychological literature, we cite Honorton and Farrari's meta-analysis of the forced-choice experiments from 1935 to 1987 as convincing evidence for the phenomenon.[1] Using 309 studies, they found overall significance ($Z = 11.41, p \leq 1.8 \times 10^{-30}$). In their analysis, they examined a variety of variables, including file-drawer and quality considerations.

Given that precognition is real, it appears that we have little understanding about its physical process. One approach to understanding precognition, has been to examine quantum mechanics. De Beauregard has explored the inherent time symmetries at the microscopic scale and the correlations implied by the Einstein-Podolsky-Rosen paradox. He found that precognition is consistent with fully relativistic quantum mechanics.[2,3] Schmidt has suggested that retro-causal interactions (i.e., PK effects propagating backward in time) is also not inconsistent with quantum theory.[4,5]

More recently, E. Targ and R. Targ conducted an experiment to explore the structure of precognition.[6] Namely, do individuals have access to actualized or probable futures? Their evidence suggests that precognition is independent of *a priori* target probabilities. However, in a similar experiment, Radin found significant evidence contradicting their result.[7]

In a carefully constructed experiment, Vassy found that a goal-oriented model of precognition did not fit his data.[8] His result was at least superficially inconsistent with Schmidt's finding that RNG results were *independent* of the internal complexity of the hardware.[9]

One of us (May) has been exploring the nature of precognition and suggests that precognition experiments are relatively easy to construct and nearly impossible to understand. For example, *if* precognition is goal-oriented (i.e., individuals can "peek" into the answer book), then process-oriented experiments are difficult, if not impossible, to interpret. It is always possible to include *all* the complexity of any given experiment into a "black box" and consider the final result as the "answer book." Thus the intervening complexities are simply not apropos.

One obvious problem with this perspective is that goal-oriented precognition is *not* falsifiable—an unacceptable circumstance in science. For any experiment to be valid, there must be a result. No matter how

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

II METHOD OF APPROACH

1. Conceptual Description

In a typical remote viewing (RV) trial, a viewer and monitor are sequestered in a laboratory. An assistant randomly selects the target from a pre-defined set (i.e., target pool). At a pre-determined time, the viewer attempts to describe the target. At the end of the session, the data are secured, and the intended target is shown to the viewer as feedback. Normally the feedback includes a presentation of the target and a complete debriefing of the RV experience.

In this experiment, we eliminated all debriefing of the RV experience, and presented the feedback tachistoscopically. The display intensities varied from zero to a level that just exceeded the visual recognition threshold (i.e., that intensity where the target could be easily seen and described in detail). Extreme care was taken in order to insure that the viewer was the only individual who was simultaneously aware of both the target and the response.

Figure 1 shows four hypothesized feedback dependencies:*

(1) Precognition—The information originates from the future feedback experience.
(2) Real-time—The information originates from the target during the remote viewing.
(3) Mixture—At low feedback intensities, the information originates from the target, but at higher intensities, it originates from the feedback.
(4) Null Hypothesis—No evidence of remote viewing is observed.

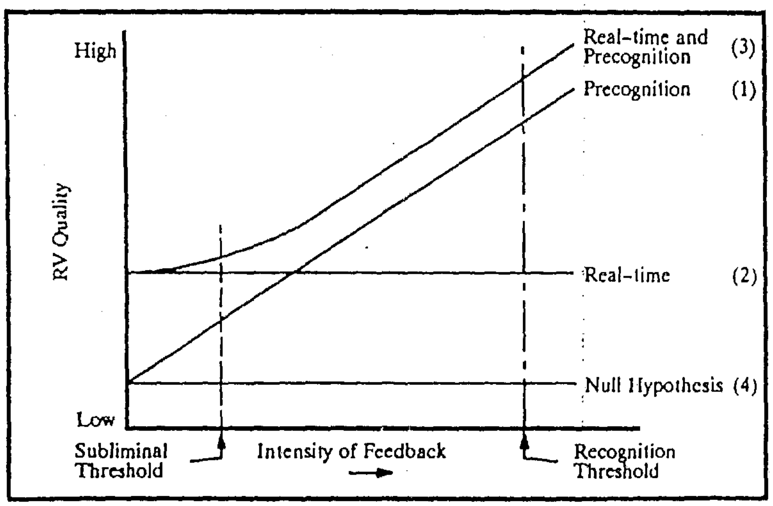


Figure 1 Idealized Relationships Between RV Quality and Intensity of Feedback

* Assuming, of course, that precognition is not exclusively goal oriented.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

One important implicit assumption must be true before the various hypothetical models shown in Figure 1 can be valid. Namely, the feedback experience is assumed to be proportional to the *cognitive* awareness of the visual feedback. Based on this assumption, the amount of information available as feedback constitutes the independent variable.

2. Detailed Description

2.1 Calibration of the Tachistoscopic Intensities

We assume that the magnitude of the feedback is directly proportional to the duration of the viewer's exposure for a given level of luminance, or (because of Bloch's Law) proportional to the luminance at a fixed exposure time.* A tachistoscope was loaded with 80 color photographic 35 mm slides (5 opaque and 75 having various luminance contrasts) of natural and man-made scenes (photographs from *National Geographic*) randomly chosen from a larger pool of 200.†

Figure 2 shows an ideal relationship between visual detectability and luminance for this experiment. To lessen the sensitivity to individual differences in perceiving the feedback, we required that the psychometric curve rise rapidly through the 50% detection level. We identified eight levels of feedback intensity that would be used as the values of the independent variable in the experiment. Two were below subliminal threshold (i.e., 0.0 detection probability), and one of these was completely opaque. Three were at 25, 50, and 75 percent detection threshold, respectively, and two were above recognition threshold (i.e., > 100 percent). These points are shown as arrows in Figure 2.

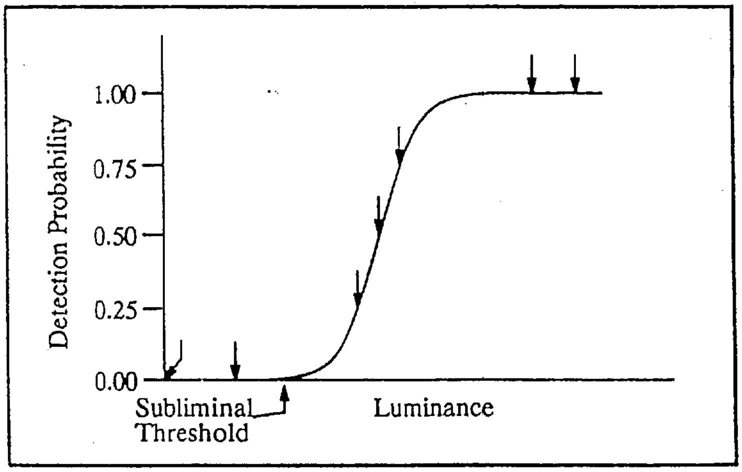


Figure 2 Ideal Psychometric Function

Visual detection, however, is not related to luminance alone. For a fixed luminance, scenes with differing contrasts will be detected with differing probabilities. At the same luminance, for example, a photograph of a checkerboard will be easier to detect than a photograph of a mountain cabin in a snow storm. Thus, each of the candidate target slides had to be calibrated with human observers in order to determine the empirical relationship between detectability and luminance.

* For presentation times shorter than about 100ms, the product of time and intensity is a constant.[10]

† The opaque slides were also photographs from the *National Geographic* magazine, but they were covered with opaque black tape.

To accomplish this, we varied the luminance contrast of the slides over an intensity range of 2^{12} by duplicating them at one of 12 f-stops (including 0 luminance).

The 80 target slides were back-projected by a Gerbrands G1170 two-field projection tachistoscope onto a 14-inch-square frosted glass window. The tachistoscope was programmed to present each slide in numerical order for 50 ms, followed by a five-second pause during which the next slide was cycled into position. Slides were attenuated by projecting them through a pair of plane polarizers: one fixed and the other variable. The luminance of the projected image could be adjusted by rotating one of the polarizers.

Two females participated in the calibration.* A complete data set was obtained from one, and data trends were confirmed by the other.

The calibration procedures were as follows. The subject was seated approximately 1 m from the projection screen, which was positioned at eye level in the wall between the room which housed the apparatus and the experiment room. The subject was permitted to view the screen and the other contents of the room freely for several minutes to ensure that she adapted to the ambient illumination level. To mask the sounds of the tachistoscope, the subject listened to white noise through earphones. The response was registered by a foot switch that the subject pressed to indicate detection of the target.

In a typical calibration session, the variable polarizer was set at a predetermined value, and each of the 80 slides was presented five times. Two sessions were conducted at each polarizer setting, providing 10 data points per slide per polarizer setting. An alternative procedure was used when the variable polarizer was set near one of the extremes of the experimental range. (By "extreme" we mean that the subject saw nearly all slides or very few of them.) To reduce the tedium, only those slides near the detection threshold were presented.

Figure 3 shows a sample of the psychometric curves generated from these data. Six of the 80 slides are shown by plotting the probability of detecting a given slide as a function of the variable polarizer setting. The following example illustrates the procedure that was used for all target slides to select those that met the criteria shown in Figure 2. Using the data shown in Figure 3:

(1) Fix the variable polarizer at 50 degrees.
(2) Observe that slide *s1* is detected with 0.5 probability (i.e., one of the requirements shown in Figure 2).

Suppose that we wish to include slide *s6* in the 0.5 detection group. It was detected, however, 50 percent of the time when the polarizer was set at 87.5 degrees—a change of 37.5 degrees from 50 to 87.5 degrees. To account for this difference:

(3) Compute the cosine of the angle difference and square it [i.e., $\cos^2(37.5) = 0.629$].
(4) Decrease the exposure of slide *s6* until the luminance is reduced by 0.629.

Now slide *s6* will also be detected only 50 percent of the time when the polarizer is set at 50 degrees.

In actual practice, we could only change exposure by an integral number of f-stops, so changes of intensity were always by factors of 2. Working backward, however, the position of all the target slides on the curve shown in Figure 2 could be determined. Furthermore, all the slides were grouped into bands near the arrows shown in Figure 2.

* They did not participate in the remote viewing portion of the experiment.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

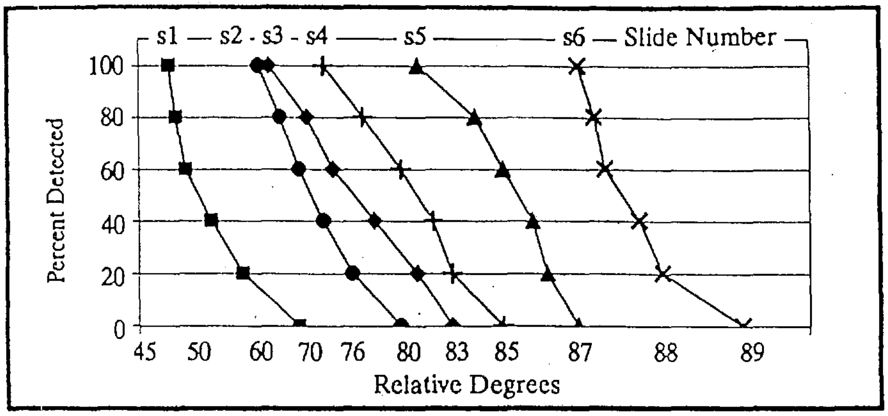


Figure 3 Degrees of Polarizer Rotation (Scaled for Equal Luminance Intervals)

Normally, data would be collected from a larger sample of individuals in order to arrive at an average function, but in this experiment, data from two persons were sufficient for several reasons. First, pilot studies indicated that interpersonal variability of target slide detection was quite low. Second, to collapse interpersonal variability even further, we generated a steep psychometric curve by sampling the abscissae coarsely. For example, if we sampled target slide contrast at only two values—0 and 100 percent contrast—all observers would respond identically, thus eliminating interpersonal variation. In this study, we sampled target contrasts at intervals that were found in pilot studies to produce low interpersonal variability. Finally, for the purposes of this study, interpersonal variability was not significant because it only shifts the psychometric function along the abscissa by some unknown amount without changing the shape of the function. Thus, interpersonal variability could only result in an erroneous estimate of the absolute magnitude of the feedback, but the relative magnitude of the feedback is independent of these errors.

2.2 Experiment Protocol

Forty targets (selected from the original 80) were prepared into eight intensity groups of five targets each using the calibration data described above. Each intensity group represented the cognitive awareness that each viewer would experience (on the average) from the feedback. The top two intensities were sufficient to experience nearly complete cognitive awareness of the feedback. By definition, those below subliminal threshold could not be cognitively sensed.

To attempt to maintain some control over precognitively available "answers," we arranged that at no future time would a response be cognitively compared to its intended target. Three pieces of information are needed to provide complete knowledge of a session: (1) the target, (2) the response, and (3) the comparison between them. The target system was prepared by individuals who had no access to the responses. Furthermore, the RV monitor, the assistant, and the viewers had no access to the targets. Finally, the analysts were never informed which were the correct results on a trial-by-trial basis.

The slide tray in the tachistoscope was controlled by a computer (Sun Microsystem 3-160) in such a way that everyone was blind to target selection during a trial. To avoid cuing, for example, the tray always began and ended in the zero position. When the computer moved the tray, an independent electrical unit, which could be accessed by the computer, counted the tray steps to assure that the intended target was displayed at the correct time.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Three experienced viewers (Viewers 009, 105, and 177) and one novice (Viewer 137) each contributed 40 trials (five at each of the eight intensity levels). Each viewer averaged about 5 trials per week.

A random order of intensities of feedback was determined (by computer) once (and differently) for each viewer prior to the start of the viewer's first trial. Once the order had been set, the trials cycled through the list of intensities until the 40 trials were complete. The sequence of events for each trial was as follows:

(1) A monitor and a viewer entered a laboratory that contained a table, two chairs, a computer terminal, and a covered 14-inch-square frosted glass window. The window served as a projection screen for the tachistoscope in the adjacent laboratory.

(2) When the viewer was ready for the session, the monitor initiated an automatic target selection program on the terminal.

(3) Using a standard feedback shift-register algorithm, which was seeded from the system clock, the computer randomly selected (with replacement) a target from within the set of five for the given intensity, stepped the slide tray to that target, and notified the monitor that the trial could begin. Because of the closed tachistoscope shutters, no illumination of the slide was present on the frosted screen.

(4) The monitor indicated that the trial should begin. For the next 10–15 minutes, the viewer drew and/or wrote responses to the intended target.

(5) At the conclusion of the session, the monitor collected the response, and the viewer opened the screen cover in such a way to shield the monitor from the feedback material.

(6) When the viewer was ready, he or she pressed a button that initiated a single tachistoscope display of the target. One, and only one, 50–ms display appeared on the translucent window screen. (Electronics prevented the viewer from receiving more feedback after the first button press.) The monitor was instructed *not* to discuss the experience with the viewer in any way at any time.

(7) The monitor ended the session, and notified the control program from the computer terminal. After the computer had returned the slide tray to zero, then, and only then, did the monitor and viewer leave the room. All target data were preserved in a computer file.

2.3 Data Analysis

The rank-order analysis used in this experiment has been described elsewhere, so only an overview is presented here.[11] Using cluster analysis, all 200 targets had previously been assigned to orthogonal clusters of similar targets (i.e., every cluster of similar targets differed from every other cluster). An assistant prepared packages (one for each viewer) consisting of all the responses randomly ordered. Next, the assistant computer-generated a list (ordered by target number) of seven targets for each response consisting of the actual target and six decoys (a different set of seven for each response). The decoys were chosen from clusters different from each other and different from the target cluster. The decoy clusters were chosen randomly from a set of 18, weighted by the number of targets in each cluster. Once a cluster was selected, the decoy was randomly selected from within the cluster.

The response material and the target/decoys set of seven photographs (i.e., one target, six decoys) were presented to two analysts for judging. The analysts arrived at a consensus to rank order each set of seven targets for each response in accordance with the best to the worst response/target match. For each viewer, a sum-of-ranks statistic was computed for the sessions. In addition, the data were plotted as RV quality (i.e, one minus the assigned rank) versus feedback intensity.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

III RESULTS AND DISCUSSION

Table 1 shows the sum of ranks, associated p-value, and effect size for each viewer in this experiment.

Table 1

Results of the Tachistoscope Feedback Experiment

Viewer	Results		
	Sum of Ranks	p-value	Effect Size (r)
009	131	0.012	0.357
105	182	0.962	−0.281
137	159	0.484	0.006
177	104	3.5×10^{-6}	0.711

Viewers 009 and 177 produced independently significant results (1-tailed). We can combine the data for all viewers in many ways, but the most conservative is a binomial calculation assuming an event probability of 0.05. Two successes in four trials corresponds to an exact p-value of 0.014. A more realistic estimate is provided by a minimum p-value technique[12] which yields 1.4×10^{-4}. The important point, however, is that this experiment produced strong evidence for an information-transfer anomaly.*

Figures 4 through 7 show RV quality plotted against intensity of the feedback for the four viewers. The lowest possible quality was zero and the maximum was seven. Shown also is a regression line and its associated linear correlation coefficient for each viewer. These figures should be compared to Figure 1, the hypothetical expectations.

The relationship that is easiest to understand is hypothesis 1 in Figure 1 (i.e., increased RV performance with increased feedback intensity). We did not observe any such correlation with either of the significant viewers. In fact, the linear correlation coefficients were not significantly different from zero.

This lack of positive correlation, in conjunction with significant evidence of RV, complicates the interpretation considerably. The most obvious conclusion is that the viewers obtained their data in real time and not from the feedback. However, because of the conceptual difficulties outlined in the introduction of this paper, alternative explanations must be considered.

* As an aside, we note that all viewers complained about the lack of "experiential" feedback.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

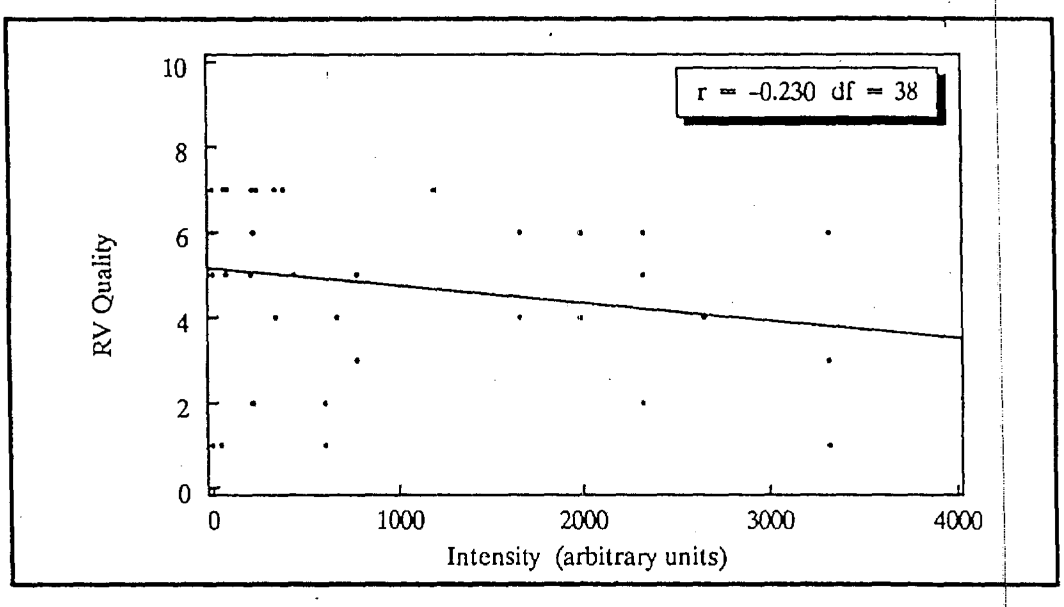


Figure 4 RV Quality as a Function of Feedback Intensity: Viewer 009

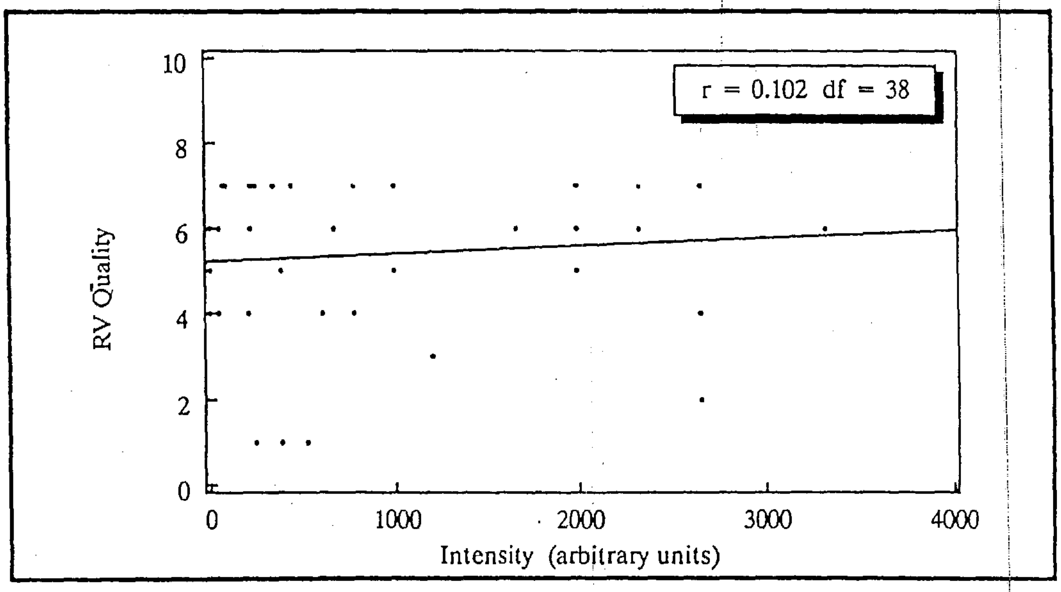


Figure 5 RV Quality as a Function of Feedback Intensity: Viewer 177

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

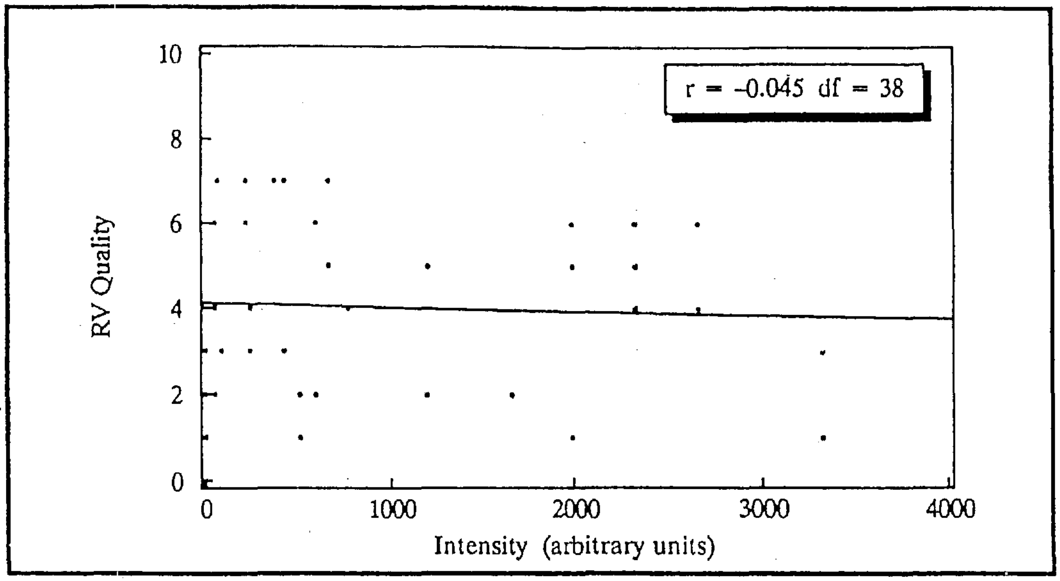


Figure 6 RV Quality as a Function of Feedback Intensity: Viewer 137

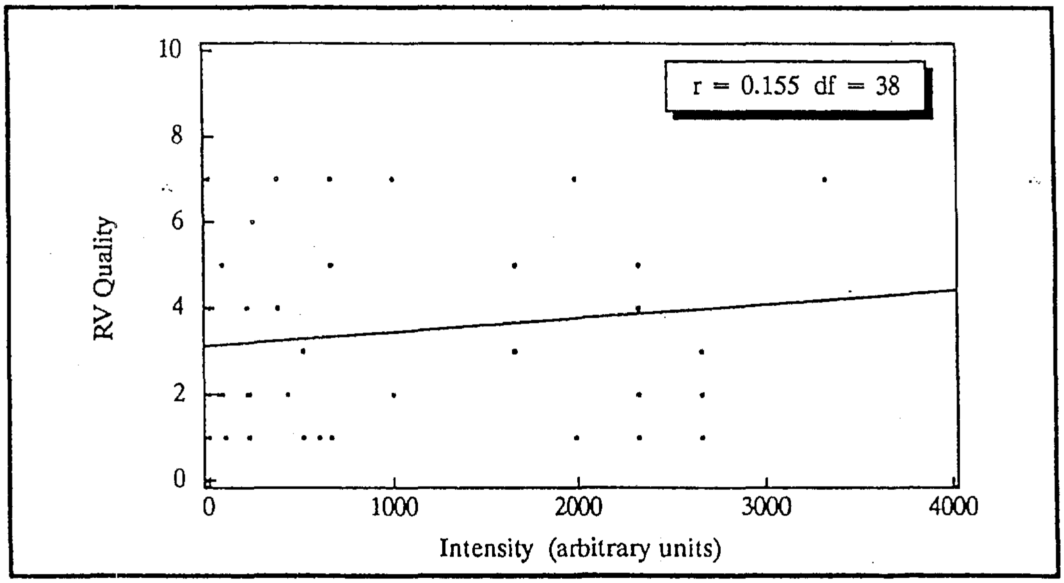


Figure 7 RV Quality as a Function of Feedback Intensity: Viewer 105

One alternative is that the absolute position of the psychometric curve is important. Feedback might be related to the cognitive experience but only at large values of luminance contrast. If this is true, then we might not expect functional dependence of remote viewing quality on the feedback in this particular experiment.

The most important alternative, however, is that precognition may be exclusively goal oriented, and thus, we are faced with the unfalsifiablity issue. We might not ever be able to interpret process-oriented experiments if this aspect of precognition is true. Therefore, the question of from what time frame does RV data originate remains unanswered at the present.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

ACKNOWLEDGMENTS

We would like to thank Beverly Humphrey for her valuable contribution as an analyst and Thane Frivold for creating the tricky computer control code. Thanks also to the tireless effort of the remainder of the Cognitive Sciences Program for target selection and technical administration.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

REFERENCES

1. Honorton, C. and Ferrari, D. C., "'Future Telling:' A Meta-analysis of Forced-choice Precognition Experiments, 1935-1987," *Journal of Parapsychology,* Vol. 53 (December 1989).
2. De Beauregard, O. C., "Time Symmetry and the Einstein Paradox," *Il Nuovo Cimento,* Vol. 42B. No. 1 (November 1977).
3. De Beauregard, O. C., "On Some Frequent but Controversial Statements Concerning the Einstein--Podolsky—Rosen Correlations," *Foundations of Physics,* Vol. 15, No. 8 (August 1985).
4. Schmidt, H., "A Logically Consistent Model of a World with Psi Interaction," *Proceedings of an International Conference on Quantum Physics and Parapsychology (1974),* pp. 205-228, Parapsychology Foundation, Inc., New York (1975).
5. Schmidt, H., "Toward a Mathematical Theory of Psi," *The Journal of the American Society for Psychical Research,* Vol. 69, No. 4, pp. 301-320 (October 1975).
6. Targ, E. and Targ, R., "Paranormal Perception as a Function of Target Probabilities," *Journal of Parapsychology,* Vol. 50, No. 1, pp.17-28 (March 1986).
7. Radin, D. I., "Effects of a Priori Probability on Psi Perception: Does Precognition Predict Actual or Probable Futures?" *Journal of Parapsychology,* Vol. 52, No. 3, pp. 187-212 (September 1988).
8. Vassy, Z., "Complexity Dependence in Precognition," *Journal of Parapsychology,* Vol. 50, No. 3, pp. 235-270 (September 1986).
9. Schmidt, H., "Comparison of PK Action on Two Different Random Number Generators," *Journal of Parapsychology,* Vol. 38, No. 1, pp. 47-55 (March 1974).
10. Marks, L. E., *Sensory Processes: The New Psychophysics,* Academic Press, New York and London (1975).
11. Humphrey, B. S., E. C. May, and J. M. Utts, "Fuzzy Set Technology in the Analysis of Remote Viewing," *Proceedings of the Parapsychological Association 31st Annual Convention,* pp. 378-394, Montreal, Canada (August 1988).
12. Hedges, L. V. and J. Olkin, *Statistical Methods for Meta-Analysis,* p. 34, Academic Press, Orlando, Florida (1985).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

V MAGNETOENCEPHALOGRAPHY

The papers in this section primarily support the anomalous cognition neuromagnetic experiments.

The number that appears in the upper right–hand corner of the first page for each publication is keyed to the following descriptions:

22. Biotechnologies, Inc., "Identification of Sources of Brain Activity Using Magnetic Source Localization and Magnetic Resonance Imaging," Research Review. Biotechnologies, Inc., publishes a number of very brief research reviews. This one discusses the potential for localizing the response to a 1000 Hz audio stimulus.
23. Sato, S., Balish, M., and Muratore, R., "Principles of Magnetoencephalography," *Journal of Clinical Neurophysiology,* Vol. 8, No. 2, pp. D1-D13, (1991). This paper is a technical tutorial for magnetoencephalography. Although it is clinically oriented, the discussion is broad enough to cover matters of interest to the research community (e.g., evoked responses).
24. Lehmann, D., Ozaki, H., and Pal, I., "EEG Alpha Map Series: Brain Micro-states by Space-oriented Adaptive Segmentation," *Electroencephalography and Clinical Neurophysiology,* Vol. 67, pp. 271-288, (1987). This technical research paper is included because it is a primary reference for paper number 19. Among other technical issues, it describes the phase shift of alpha activity as a result of external stimuli.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

RESEARCH REVIEW

Identification Of Sources Of Brain Activity Using Magnetic Source Localization And Magnetic Resonance Imaging

Pantev, C., Hoke, M., Lehnertz, K., Lutkenhoner, B., Fahrendorf, G., Stober, U. Identification of sources of brain neuronal activity with high spatiotemporal resolution through combination of neuromagnetic source localization (NMSL) and magnetic resonance imaging (MRI). *Electroencephalography and clinical Neurophysiology,* 1990; *75:* 173-184.

In this study, the locations of the origin of the M100 wave of the auditory evoked magnetic field in response to tone bursts of different frequencies, obtained through dipole localization methods, were related to cerebral structures by comparison with coronal MRI tomograms of each subject.

Four normal subjects were presented with contralateral tone burst stimuli of 250, 500, 1000, 2000, and 4000 Hz, and average evoked magnetic fields were recorded over a matrix of positions covering one hemisphere. The waveform data for one subject at 60 recording positions are shown in Figure 1, as is an equi-contour map of the magnetic field values at the instant of time where the M100 peak reaches its maximum value for the 1000 Hz stimulus.

The dipole localizations for the M100 peak for stimulation with tone frequencies from 250 Hz to 4000 Hz confirmed previous results showing a linear increase in the depth of the dipole source with the logarithm of tone frequency. Moreover, all source localizations corresponded to the primary auditory cortex located on the surface of Heschl's gyrus. A comparison of the dipole localizations for one subject with the subject's MRI tomogram is shown in Figure 2.

The results highlight the exceptional capabilities of a combination of these two non-invasive, high-resolution techniques for functional diagnosis.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

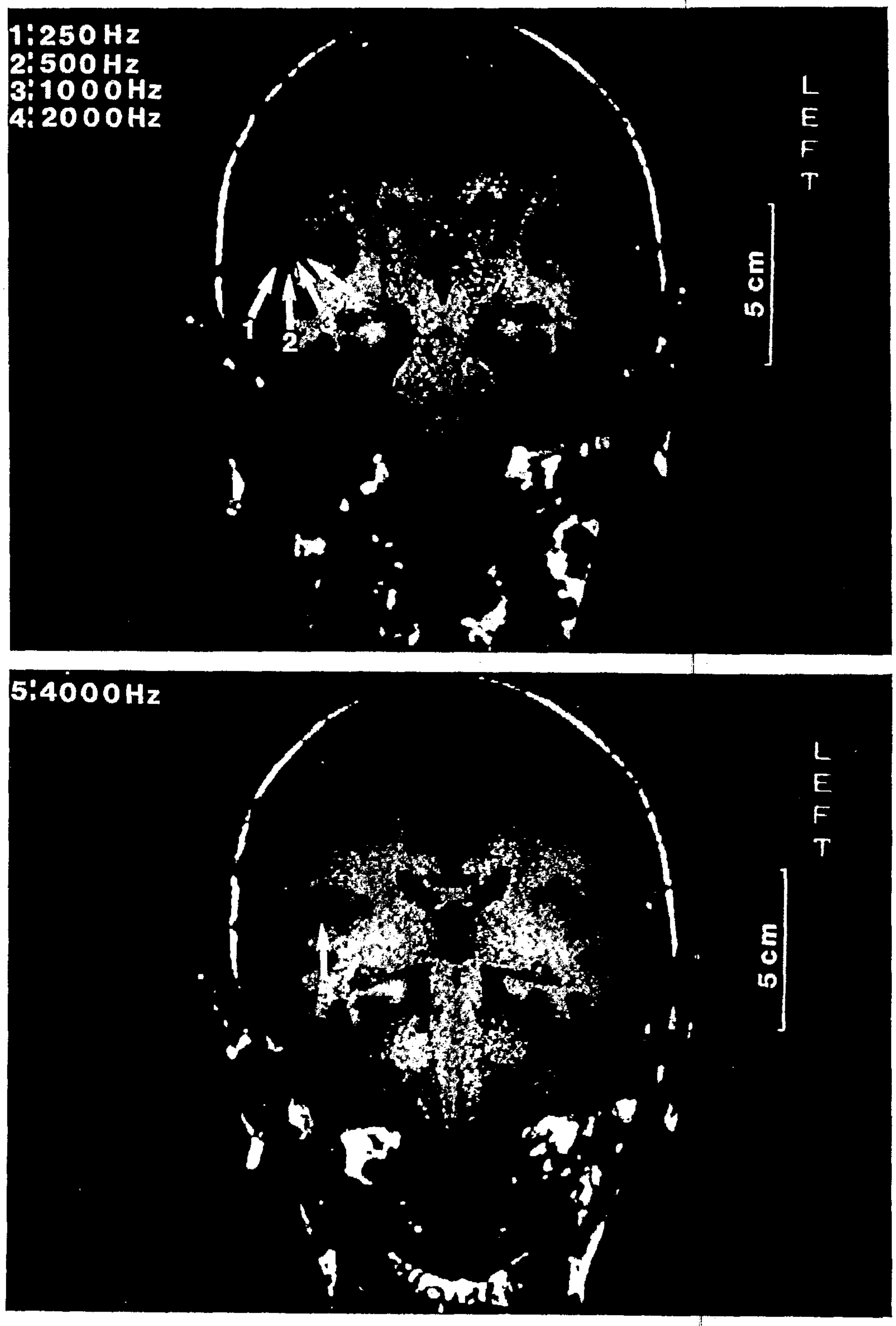


Fig. 2. Comparison of the results of functional localization obtained for one subject (Fig. 2) with anatomical structures obtained from the same subject through magnetic resonance imaging. ECD locations are marked by black dots at the tips of white arrows

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

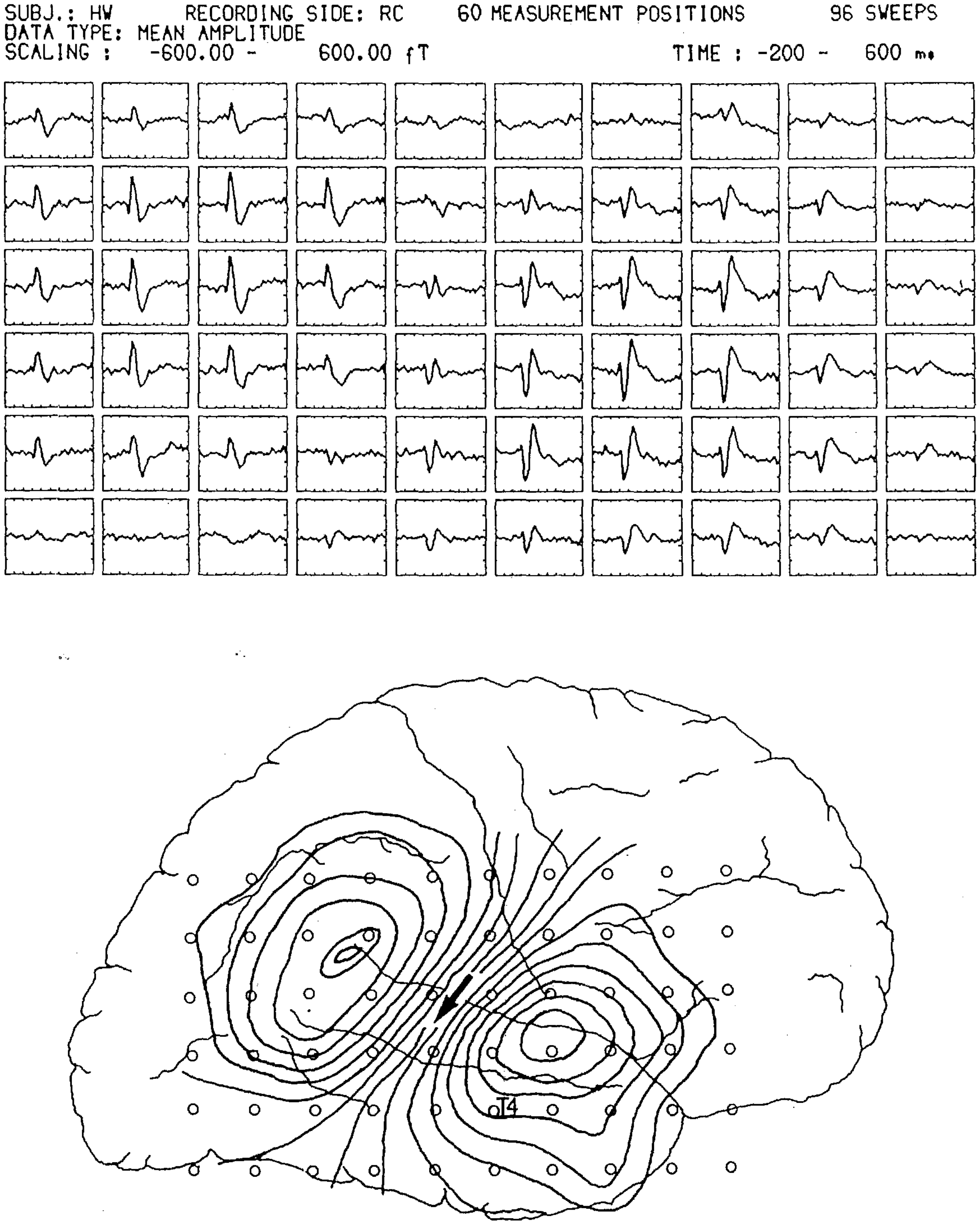


Fig. 1. Top: synopsis of averaged response wave forms measured from 60 positions over the right cerebral hemisphere (stimulus frequency 1000 Hz). Bottom: distribution of the auditory evoked field at that instant of time when wave M100 assumes its maximum, projected onto the rough outline of the supralateral surface of the brain. Circles: measurement positions; arrow: calculated position and direction of the ECD (taken from Hoke 1988).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Journal of Clinical Neurophysiology
8(2):D1-D13, Raven Press, Ltd., New York
Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4
© 1991 American Electroencephalographic Society

Principles of Magnetoencephalography

Susumu Sato, Marshall Balish, and Robert Muratore

Neurophysiology Unit, Medical Neurology Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland, U.S.A.

Summary: Magnetoencephalography (MEG) is a new, noninvasive functional test equivalent to EEG. It has been used to localize the sources of evoked responses and interictal and ictal epileptiform discharges and to study patients with psychiatric illnesses, cerebrovascular accidents, and migraine. In epilepsy research, it is hoped that MEG will provide information similar to that yielded by depth or subdural electrode recording, or that the combination of these methods will provide more information than either one alone. The application of MEG appears to be widening, although it is not yet a routine clinical diagnostic tool. The utility of MEG is limited by technological problems, but new and more efficient systems are becoming available. Within several years, advances in the technology and understanding of MEG may modify the course of its application.
Key Words: Magnetoencephalography—Instrumentation—EEG—Evoked responses—Epilepsy.

Magnetoencephalography (MEG) is a new, noninvasive functional test equivalent to EEG. MEG has been used to localize the sources of evoked responses and interictal and ictal epileptiform discharges and to study patients with psychiatric illnesses, cerebrovascular accidents, and migraine. In epilepsy research, it is hoped that MEG will provide information similar to that yielded by depth or subdural electrode recording, or that the combination of these methods will provide more information than either one alone (Wood et al., 1985).

PHYSICS OF MEG

Generation of Biomagnetic Fields

The decrease in force with increasing distance of both the magnetic and electric forces suggests that electricity and magnetism are related. The relation between a magnetic field and its associated current was empirically defined by Jean Patiste Biot and Felix Savart. Biot-Savart's law states that the magnetic field due to a small current element varies as the inverse square of the distance of the current element, varies directly with the current, and varies with the sine of the angle between the directions of the current and of the [vector] leading to it. In modern notation,

$$dB = \frac{\mu_0}{4\pi}\frac{i d\underset{\sim}{l} \times \hat{r}}{r^2}$$

where dB is the differential magnetic field in tesla, μ_0 is the magnetic permeability of free space and is assigned the value of $4\pi \times 10^{-7}$ Tm/A, i is the current in amperes, $d\underset{\sim}{l}$ the length of the current element in meters, $\hat{r}$ is the unit vector pointing from the location in space at which the magnetic field is evaluated to the location of the current element, and r is the distance between those two locations in meters. The cross-product can be interpreted by the right hand rule: dB is mutually perpendicular to both $d\underset{\sim}{l}$ and $\hat{r}$. This relation is understood by placing the right hand in space at the location at which the magnetic field is

Address correspondence and reprint requests to Dr. S. Sato at Building 10, Room 5C101, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, U.S.A.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

to be evaluated. Point the hand parallel to the current element *idl*. Excepting the thumb, point the fingers parallel to $\tilde{r}$. Extend the thumb, and it will point along the direction of *d*B. A bit of finger play shows that a magnetic field encircles a straight wire carrying current and forms a dipolar (donut-shaped) region around a loop of current. The superposition principle, which states that the field produced by a sum of sources is equal to the sum of the fields of the individual sources, allows the total magnetic field **B** to be calculated by summing up the contributions from the individual current elements. In this way, the magnetic field for a particular source model can be determined. Williamson and Kaufman (1990) have reviewed this subject in detail.

Instrumentation

The measurability of the neuromagnetic field produced by the flow of electric current in cerebral neurons is made more likely by the columnar arrangement of the cortical neurons, which leads to the summation and re-enforcement of the field of each of the firing neurons. Nonetheless, a typical neuromagnetic field is on the order of 10^{-12} T, several orders of magnitude less than the magnetic field of the earth. The measurability of the magnetic field is made successful by shielding the subject and the measuring instrument from extrinsic noise, and by the exquisite sensitivity of the measuring instruments.

Measuring Devices

If an electric current can give rise to a magnetic field, it is natural to ask if a magnetic field can give rise to an electric current. A magnetic field can give rise to an electric current as long as the magnetic field is changing. This effect is called induction and is the basis for the detection of magnetic fields. If a coil of wire is placed in a changing magnetic field, an electric current is generated in the coil. This current then becomes the signal of the field. Reorienting the coil and changing the number and distribution of wire turns in the coil change the signal.

One device has become standard in neuromagnetic measurements, the superconducting quantum interference device (SQUID), which allows very sensitive measurements of the magnetic field. A SQUID is made of a ~~metal~~ niobium ring with a weak link, an extremely thin gap, across which tunneling occurs.

There are two types of SQUID: rf-SQUID, which has a weak link and is biased with radiofrequency current, and DC-SQUID, which has two weak links and is biased by DC current. The latter is less noisy than the former and therefore has been used in most of the recent magnetometers. The SQUID itself can be used as a magnetometer, but its combination with detection coils of various configurations allows a flexible design of measuring devices.

Erne (1983) has reviewed SQUIDs as used in biomagnetism.

Detection coil configurations can be classified by the aspect of the magnetic field to which they are most responsive. In this system, a magnetometer is a zeroth-order gradiometer. A first-order gradiometer is sensitive to the first spatial derivative of the field, and so on. Coils aligned axially are sensitive to the derivatives along their axis, and coils aligned in a plane are sensitive to the derivatives in that plane. The most common design is a second-order axial gradiometer with its axis approximately perpendicular to the head. However, this configuration introduces the possibility of signal loss or reduction of sensitivity by virtue of layers of coils (the distance between adjacent coils is a baseline). The recent trend is to utilize the first-order gradiometer configuration. In fact, the first-order gradiometer is already used in the recent larger neuromagnetometers (Biomagnetic Technologies, 1989; Gudden et al. 1989).

In a typical arrangement, detection coils are inductively coupled to a SQUID. Associated circuitry provides an electrical potential that is proportional to the magnetic field near the detection coils. The detection coils and SQUID are cooled to superconducting temperatures by a liquid helium bath in a dewar. Although its primary purpose is the operation of the SQUID, the cold bath provides some benefits and inconveniences. The chief incidental advantage of the cold bath is the great reduction in noise. Among the inconveniences are the need to maintain the liquid helium and the necessity for dewar insulation, which inhibits the mechanical positioning of the device ~~beyond a 45° tilt~~. These problems hamper efforts to incorporate large numbers of sensors into the dewar. A cryocooler system with multiple stages of gaseous cooling was developed to try to overcome this inconvenience. The prototype was used at New York University, with some success; however the number of SQUID coils is limited and there are noise sources introduced.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Increasing the number of sensors in the dewar provides the ability to take simultaneous recordings across large areas of the head, as in EEG. Early MEG systems used a single channel. As the arrays grew, a hexagonal arrangement became popular, and the next largest size was the 7-channel machine. Currently, 2 more rings of coils have been successfully added, producing 37-channel machines that cover an area of 14.4 or 20 cm in diameter for the Biomagnetic Technologies system and Siemens' system, respectively. Machines with as many as 100 channels for full head coverage are being designed.

The use of high-temperature superconducting material is well in line with technology advancement in biomagnetism but preliminary experiments suggest that a SQUID made with such materials is too noisy to be of practical use (Clarke and Koch, 1988).

Shielding

Magnetic shielding is typically provided by building a smaller room inside a larger one. The smaller room houses the dewar with the SQUID, associated circuitry, a gantry for movement of the dewar, and a chair or bed for a human subject. The inner room is called the shielded room, and its walls are usually made of several layers of a conductor such as aluminum or copper, which provides eddy current shielding of higher frequency noise, and several layers of mu-metal or other highly magnetically permeable material, which provide shielding of lower frequency noise. Erne and Romani (1990) have reviewed magnetically shielded rooms.

In some measuring environments, gradiometer coil designs provide spatial filtering of magnetic fields, which might obviate the need for extensive and expensive shielding. In the hospital environment, however, with many medical machines and numerous other sources of interference, a magnetically shielded room is mandatory.

COMPARISON OF MEG AND EEG

There are some important differences between EEG and MEG. The electrical potentials measured by EEG are field potentials, or so-called volume cur-

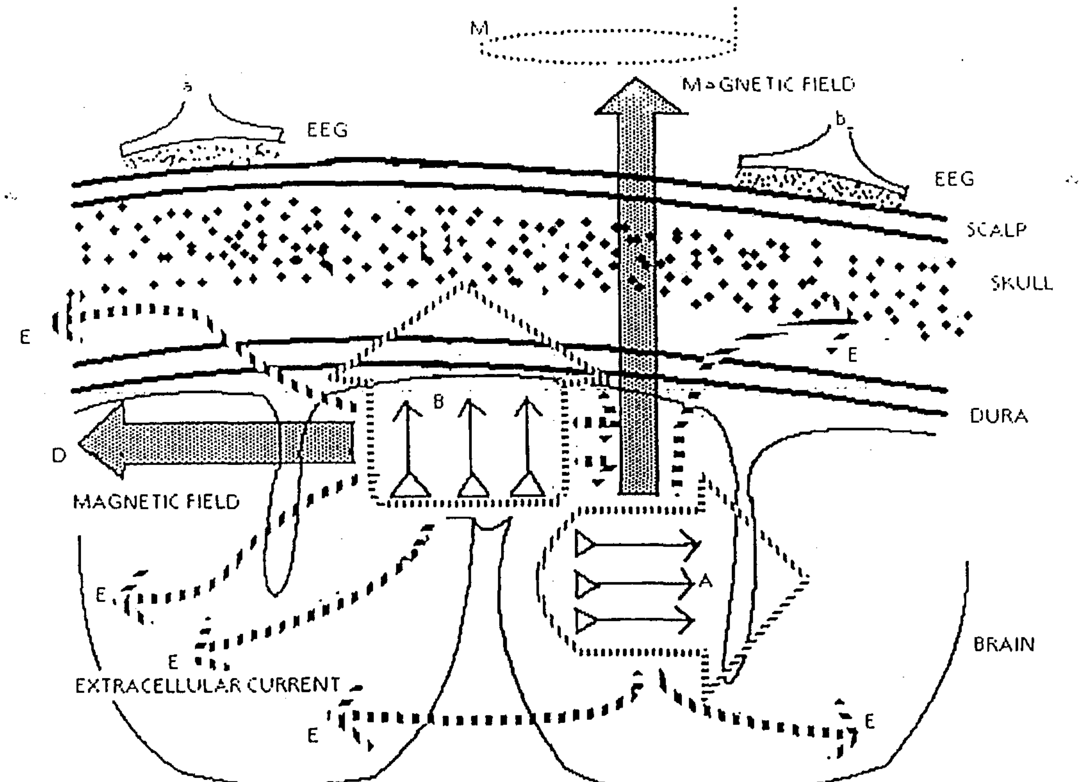


FIG. 1. EEG is recorded between two electrodes (a and b) placed over the scalp, whereas MEG records the magnetic fields emanating from a tangential source (A). The magnetic field (C) is not distorted by the scalp, skull, and dura, but the EEG signal is dispersed, attenuated, and distorted by them (E). A radial dipole (B) generates magnetic fields (D) parallel to the detection coil (M), and they may not be detected.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

rents; they are easily attenuated, diverted, and distorted by the dura, skull, and scalp, whereas the magnetic fields are not affected by these structures (Fig. 1). However, it is well known that the magnitude of the magnetic fields decays at a rate of $1/r^2$ to $1/r^3$ (where r = distance from the source). Therefore, there is some concern about the ability of MEG to detect deeper sources (Cuffin and Cohen, 1977*a*). Electrodes and conductive gel are used in EEG, but not in MEG. The quantity measured by EEG is a relative value, namely, a potential difference between two electrode positions, whereas MEG measures an absolute value when the distance of a source is within the baseline of the detection coil (Wikswo and Roth, 1988) and a relative value when a source is deeper than the baseline distance. The EEG contour maps are often monopolar and occasionally dipolar. The MEG maps are usually dipolar or multipolar and rarely monopolar. The MEG pattern is rotated by 90° from the EEG pattern, and the MEG maps are about one-third tighter that the EEG maps (Cohen and Cuffin, 1983). EEG is known to measure the field potentials of current sources oriented in all directions, with radially oriented current sources measured somewhat preferentially. MEG measures chiefly the radial component of the magnetic fields, because of the rough sphericity of the head and the configuration of the magnetometer. This peculiarity of MEG has been interpreted to mean that MEG measures tangential current sources located in sulci; because about two-thirds of the cortex is in sulci, it is thought that loss of information due to MEG's inability to measure current sources in the gyri is minimal (Hari and Kaukoranta, 1985).

Calculation of three-dimensional coordinates from EEG data requires six parameters: three for location; one for magnitude; one for orientation, and one for conductivity, whereas MEG does not require a parameter for conductivity (Stok, 1986).

Because of the tissue-induced attenuation and diffusion of the EEG, the dipole depth calculated by EEG mapping may turn out to be deeper than the actual depth. The direction of current flow estimated by evoked magnetic fields is opposite the direction estimated by evoked potentials, suggesting that the magnetic fields are generated by high-density intracellular current, whereas the electrical potentials are produced by volume current (Kaufman and Williamson, 1982).

These comparisons suggest that MEG may have a somewhat better spatial resolution than EEG (Cohen and Cuffin, 1983).

MODELING IN MEG

A knowledge of the so-called forward and inverse problems is essential for understanding biomagnetism and its analysis, but there is no unique solution to the inverse problem. For MEG analysis, information concerning the geometry of the head and the current source or sources must be obtained. A simple spherical model with a single current dipole has been used most commonly, although other models may be more appropriate in some situations. These subjects have been reviewed by Nunez (1981), Ilmoniemi et al. (1985), Hari and Ilmoniemi (1986), Stok et al., (1986, 1987), Sarvas (1987), and Williamson and Kaufman (1987).

Head Geometry

The simplest model of the head is the homogeneous half-space model with a single dipole (Williamson and Kaufman, 1987), in which the magnetic field at any point is easily predicted by Biot-Savart's law, allowing estimation of the location and direction of such a dipole from the magnetic field map and the interextrema distance. In practice, the head is most commonly modeled as a uniform conducting sphere, in which radial current dipoles are assumed to have no measurable external magnetic field (Cohen and Hosaka, 1976; Cuffin and Cohen, 1977*b*). In this model, it is also assumed that volume currents make no contribution to the field and that simple relations exist between field maxima and the dipole location and depth. This is probably the most frequently used model in MEG; however, there are important problems with this model. Such a simple model does not represent the temporal lobe well, causing an angulation problem between the sensors and the inner surface of the skull, and the results are susceptible to the effects of volume currents (Barth et al., 1986; Rose et al., 1987*b*, 1989*a*).

More precise models based on the actual skull shape obtained by magnetic resonance imaging (MRI) have been suggested (Meijs et al., 1987; Meijs and Peters, 1987) and give more accurate localization in in vitro models, although at some computational expense. Such a model was recently applied to localization of scalp-recorded spike discharges in the temporal region and appeared to give reasonable accuracy of discharge localization compared with that obtained by subdural electrodes (Ducla-Soares et al., 1989).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Source Models

Modeling spike foci as a current dipole source is simplistic; the foci are probably more distributed. Nunez (1986) discussed possible effects of multiple dipole sources, and Barth et al. (1989) attempted to solve the sources of interictal activity with multiple current dipoles, each with its own temporal properties. A more complicated distributed current source is probably more representative of the epileptogenic focus, and work is underway on the interpretation of MEG sources with respect to such a model (Clarke et al., 1989; Ioannides et al., 1989; Kado et al., 1989), although this has not yet been applied to the field of epilepsy.

SPONTANEOUS ACTIVITY

Alpha Rhythm

In 1968, alpha rhythm was magnetically measured with an induction coil for the first time (Cohen, 1968). There is a good correlation between MEG and EEG alpha rhythm (Hughes et al., 1976; Cohen, 1979; Cohen and Cuffin, 1979; Modena et al., 1982). The amplitude of alpha rhythm is largest over the parieto-occipital regions (Reite et al., 1976), and the largest magnetic flux of alpha rhythm occurs over the longitudinal midline (Cohen, 1979). The maximum amplitude of the magnetic alpha rhythm is 2.5 pT peak to peak (Carelli et al., 1983). Many investigators have attempted to localize the sources of alpha activity, and some have suggested that the alpha activity has multiple sources in or near the visual cortex (Vvedensky et al., 1986; Carelli et al., 1989). Recent work suggests that magnetic alpha rhythm arises from many discrete sources, oscillating one after another and occasionally overlapping temporally. These sources are clustered near the midline, extending to a depth of several centimeters (Williamson et al., 1989).

Sleep Spindles

Sleep spindles were initially difficult to record with MEG (Hughes et al., 1976; Freedman, 1981). With a laboratory-built single-channel magnetometer, however, sleep spindles (12.5–16 Hz) were recently detected at the central vertex, predominantly as a radial magnetic component, in three normal volunteers (Nakasato et al., 1990).

Epileptiform Activity

The use of MEG in epilepsy research has been reviewed by Rose et al. (1987*c*). A worldwide surge of interest in MEG occurred when investigators from the United States and Italy published their results on localization of epileptiform discharges in 1982 and 1984 (Barth et al., 1982, 1984*a*; Modena et al., 1982; Ricci et al., 1984). MEG is used to study interictal and ictal discharges and also background activity, because all of them are important for localization of the seizure origin.

Interictal Activity

Hughes et al. (1977), followed later by Modena et al. (1982), first reported that the spike component of epileptiform discharges was well defined magnetically, but the slow-wave component was not. Simultaneous MEG and EEG recordings showed that EEG slow waves and MEG multiple spikes often occurred together or that EEG changes occurred without accompanying MEG signals or vice versa (Ricci et al., 1983; Sutherling et al., 1988*a*).

Spike averaging. At each sensor position, as many as 10–20 similar EEG spikes are sampled, and their magnetic signals are averaged using an EEG channel as a trigger. Then a map of the magnetic field is constructed and used for calculation of equivalent current dipole sources for spike discharges. The purpose of spike averaging is to improve the signal-to-noise ratio, and the eventual localization is usually expressed as a point or central point (centroid) (Sutherling et al., 1988*a*). Even though the epileptogenic focus is physiologically not a point, dipole source modeling is an important step toward further understanding of a complex problem such as epileptic seizures (Barth et al., 1982, 1984*a,b*; Sutherling et al., 1988*a*). Unfortunately, spike-averaging methods may result in loss of spatial and temporal information and of details concerning individual spike types. The template method (Salustri and Chapman, 1989) can be used to find similar spikes for averaging. With a larger MEG system, such as a 37-channel machine, this process becomes easier or unnecessary because the magnetometer does not have to be moved around to scan the patient's head.

Single-spike analysis. Individual spikes can be mapped and their location estimated without averaging (Sato et al., 1985; Rose et al., 1987*a*). With this method, important information riding on "waves" will not be "averaged out." Although individual EEG spikes appear somewhat dissimilar in morphology, amplitude, and distribution, it nonetheless seemed possible to group spikes according to their similarities, thereby yielding several different types of spikes (Rose et al., 1987*a*). With a single-channel

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

or seven-channel system, the entire head cannot be measured simultaneously, causing difficulties in combining events dispersed in time. Therefore, the selection of signals with a good signal-to-noise ratio becomes essential but is tedious. The method of identifying spike types is typically visual, although other methods are available (Salustri and Chapman, 1989). Nevertheless, less information may be lost with single-spike analysis than with averaging techniques. With a 37-channel magnetometer, consecutive spikes can be analyzed with good results on localization (Sato et al., 1990).

Relative covariance method. MEG and EEG are recorded simultaneously, the EEG activity of interest (frequency band) is identified by Fourier transform techniques, and then MEG and EEG are digitally filtered to a narrow band surrounding the frequencies of interest. The covariance of the filtered MEG and EEG signals at each sensor site is calculated and divided by the variance of the filtered EEG, giving the "relative covariance." This calculation is performed at all MEG locations, and the EEG channel or channels used are consistent for all MEG locations, although different EEG channels may be compared (Ricci et al., 1984, 1985, 1987; Romani and Leoni, 1985; Chapman, 1989). It is claimed that the relative covariance is proportional to the value of the magnetic field perpendicular to the head at a given location (Ricci et al., 1985). A contour map of the relative covariance is constructed, and an equivalent dipole current source may be localized. The localization step is typically based on a spherical model assuming that only radial magnetic components are measured (Ricci et al., 1985).

The relative covariance method is important, because it allows analysis of background activity and localization in the absence of a spike focus if there is detectable abnormal rhythmic activity. In principle, a similar analysis could be applied to stereotyped seizures with rhythmic activity.

Ictal Activity

The capturing of an ictal event with MEG is difficult, particularly with the single-channel and seven-channel systems. Nevertheless, ictal recording is not entirely precluded, because the patient usually does not move for 5–10 s during the initial phase of the seizure (Sutherling et al., 1987) or during simple partial seizures (Rose et al., 1989*b*). Simple recordings of ictal events such as 3/s spike-and-wave discharges were done early in the development of MEG (Modena et al., 1982), but detailed studies to determine spatial localization of sources are difficult inasmuch as events would have to be recorded at many locations (in the absence of large multichannel magnetometers) and the information from different seizures combined. In patients with jacksonian seizures, MEG was thought to provide localizing information despite the absence of abnormal EEG findings (Modena et al., 1982; Ricci, 1983).

Migraine

Spreading cortical depression, which is a slowly changing potential, has been implicated in migraine to explain the evolution of the clinical manifestations associated with the disorder. Although spreading cortical depression has never been observed spontaneously in humans, biphasic waves lasting for less than 10 s with amplitudes ranging from 800 fT to 13 pT were recorded with MEG in patients with migraine (Tepley et al., 1989). This area of research requires further investigation and confirmation of results.

EVOKED RESPONSES

While studies of evoked magnetic responses have not yet provided unique clinical information, they do provide information on cerebral functional processing and, when combined with imaging information, establish interesting functional–anatomical correlations. The information provided by MEG complements and in some cases resolves ambiguities remaining in comparable studies of evoked potentials. All modalities may be studied by MEG, but stimuli must not generate significant magnetic fields.

Auditory Evoked Responses

Short-Latency Evoked Responses

The detection of brainstem-evoked fields is technically quite difficult. When signals somewhat above background noise were recorded in a shielded room, using a magnetometer (not a gradiometer) and averaging from 12,000 to 96,000 sweeps (Erne et al., 1988), waveforms corresponding to brainstem-evoked potential waves V and VI could be recorded, and their amplitudes were different at different recording locations.

Middle- and Long-Latency Evoked Responses

The auditory stimulus is generally led by plastic tube to the patient to avoid interference from the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

magnetic field generated by the usual transducers or headphones. Table 1 shows recording parameters and details of stimuli from a sample of studies. Responses to approximately 100 stimuli are usually recorded at each magnetometer position. The magnetometer is moved (by some investigators randomly) over the region of interest in order to avoid systematic bias resulting from a change in state, which has a significant effect on responses. Filter settings given in Table 1 reflect the final bandpass; the signals typically are recorded through analog filters with a wider bandpass and are subsequently filtered digitally.

Reite et al. (1978) demonstrated the feasibility of recording magnetic fields to auditory stimuli. Hari et al. (1980) described the time course and spatial distribution of auditory evoked responses. In particular, major peaks at a latency of 100 ms (N100) and 200 ms (P200) were described in addition to a subsequent slow sustained field. For the most part, when source localization has been attempted, simple models such as a single equivalent dipole in half space (Pantev et al., 1988) or a single equivalent dipole in a spherical conductor (Papanicolaou et al., 1990) have been used.

Early work (Hari et al., 1980; Elberling et al., 1982) suggested that the generators of the N100 were probably located on or near the planum temporale. More recent studies with gross anatomical correlation (Pantev et al., 1988) or MRI correlation (Pantev et al., 1990; Papanicolaou et al., 1990) have largely confirmed this localization. The generators of a peak at 50 ms (Reite et al., 1988) are also near the planum temporale.

Reite et al. (1981) showed that the amplitude of the N100 peak is larger when the stimulus is from the contralateral side. Papanicolaou et al. (1990) found that the equivalent dipole for the N100 was located more posteriorly and medially when stimuli were contralateral. Rogers et al. (1989, 1990) suggested that successive time points in the N100 peak had generators that moved anteriorly, and this was clear when stimulation was contralateral, not ipsilateral. These and other studies, then, suggest the possibility that somewhat different cortical areas are activated by contralateral and ipsilateral stimulation.

Elberling et al. (1982), on the basis of the N100 response, suggested that equivalent dipoles over the left hemisphere are posterior to those over the right. Reite et al. (1988) made similar observations on the 50-ms response, although MRI correlation allowed localization of the equivalent dipole near the planum temporale.

TABLE 1. *Auditory evoked responses*

Reference	Acquisition [Freq (Hz), time (s)]	Final bandpass (Hz)	Repetitions	Number of positions	Magnetometer (channels)	Design (order)	Stimulation[a]	Shielding
Reite et al. (1978)	?, 0.2	5.0–15.0	516	5	1	2nd	Clicks; ISI, 0.25	Aluminum
Reite et al. (1981)	?, 0.5	0.1–30	128	1	1	1st	Clicks; ISI, 0.7	Aluminum
Elberling et al. (1982)	320, 0.8	DC–18	60	24	1	1st	Tones (1 kHz), 0.5; ISI, 1.5–4.5	No
Romani et al. (1982*a,b*)		32	1,000	40	1	2nd	AM tones (200–5,000 Hz) steady state	No
Pantev et al. (1988)	250, 1.024	0.1–30	96	50–60	1	2nd	Tones (250–4,000 Hz), 0.5; ISI, 4	No
Papanicolaou et al. (1990)	?	?	100	56	7	2nd	Tone (1 kHz), 0.05; ISI, 0.4	Yes
Hari et al. (1980)	?	0.03–15	>300	30	1	0	Tone (1 kHz), 0.8; ISI, 4	No

[a] Type of stimulus (frequency), length in seconds, interstimulus interval (ISI) in seconds.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Early studies using a steady-state technique (Romani et al., 1982*a,b*) suggested a tonotopic arrangement of the responses, and more recent work (Pantev et al., 1988, 1990) is in basic agreement; the depth of the equivalent dipole generating the N100 increases with the logarithm of the stimulus frequency. The mapping is not clear for the later waves at 160 ms.

Various paradigms involving speech elements as stimuli, or so-called oddball paradigms, have been used in studies attempting to elucidate cerebral processes underlying attention and speech perception (Sams et al., 1985; Makela et al., 1988; Hari, 1989*a,b*; Hari and Lounasmaa, 1989). This work is beyond the scope of the present review, but such studies, when combined with MRI findings, for example, should provide impressive functional-anatomical information.

Somatosensory Evoked Responses

The basic stimulation paradigms for somatosensory evoked responses are not greatly different from those used in recording somatosensory evoked potentials (SSEPs). Recordings may be successfully performed even without a shielded room (Sutherling et al., 1988*b*). The current pulse is brief enough not to interfere with subsequent recordings. An overview of methodology is given by Hari and Kaukoranta (1985); they also briefly discuss fields evoked by noxious stimuli. Huttunen (1986) has described methodology for recordings of fields evoked by tactile stimulation. The recording bandpass and sampling rate determine to a large extent whether or not early peaks corresponding to the "cortical" components of SSEPs are emphasized.

Several investigators have reliably recorded components corresponding to cortical SSEPs (Rossini et al., 1988; Sutherling et al., 1988*b*; Tiihonen et al., 1989). Tiihonen et al. (1989) found that the equivalent dipole for the 27-ms peak (P27m) was located on the average 1 cm anteromedially to sources for the 20-ms peak (N20m). Their findings were consistent with a tangential source (posterior bank of central sulcus) as the generator of the N20m. Sutherling et al. (1988*b*) compared localization of somatosensory evoked responses based on MEG, EEG, and electrocorticography (ECoG). The average distance of localization from the central fissure was 4 mm in EEG and MEG and 3 mm in ECoG. Averaging MEG and EEG distances yielded localizations comparable to those of ECoG. These results, although from only three patients, suggest that MEG and EEG may reliably and noninvasively localize somatosensory cortex. Suk et al. (1989) attempted to correlate localizations from stimulation of different fingers with MRI-based structural information. Their studies suggest that such noninvasive mapping of the sensory homunculus is possible.

Visual Evoked Responses

Teyler et al. (1975) were among the first to report on studies of visual evoked magnetic responses and described a method based on brief flashes. Because cathode-ray tubes generate large magnetic fields, more sophisticated stimuli (i.e., patterns or gratings) require the use of projection equipment and mirrors. Kouizjer et al. (1985) and Aine et al. (1989) described their techniques for stimulation and recording of visual evoked responses. As in other modalities, in the absence of large multiarray systems, obtaining field patterns requires repeating the averages at different locations. Filter settings and sampling are otherwise similar to recording visual evoked potentials. The appropriate treatment of field maps, even if apparently dipolar, is controversial. Stok et al. (1986) argued that the inverse problem should probably take into account a realistic head shape, rather than relying on spherical models. In spite of this objection, George et al. (1989), using sinusoidal grating stimuli, localized apparent equivalent dipoles that, when projected on MRI scans, were near calcarine cortex and appeared to evolve temporally. Under some conditions of stimulation and at some time points, two different dipoles appear necessary to explain the evoked fields (Aine et al., 1989).

Premotor Fields

Premotor fields are studied by recording magnetic activity preceding self-paced or triggered movements (Mizutani et al., 1989). Table 2 shows selected parameters used by several investigators. The electromyographic (EMG) burst is frequently used as the trigger to average the preceding magnetic activity. The procedure must be repeated at each magnetometer position, and the time transients then can be used to form contour maps. Depending on the resulting contours, an appropriate model can be used (i.e., dipole in a sphere) to make inferences about source localization. Great care must be taken to eliminate epochs with a large amount of eye or head movement (Antervo et al., 1983).

Slow changes in the magnetic field with topographic variation have been recorded as much as 1 s before foot (Antervo et al., 1983) and complex

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE 2. *Premotor fields*

Reference	Acquisition [Freq (Hz), time (s)]	Final bandpass (Hz)	Repetitions	Number of positions	Magnetometer (channels)	Design (order)	Stimulation or trigger	Shielding
Antervo et al. (1983)	175, 4	0.03–30	70	15	1	1st	EMG, foot	Yes
Cheyne et al. (1989)	?, 2	0.1–10	?	30–90	14	2nd	EMG, foot, hand, finger, face	Yes
Deecke et al. (1985)	85, 3	0.03–15	>80	11	1	3rd	EMG, fingers	No
Mizutani et al. (1989)	?, 5	0.1–20	>64	7	1	2nd	Keypress	No

hand (Deecke et al., 1985) movements. The slow-changing field preceding movement is designated the "readiness field" (RF) or "bereitschaftsmagnetfeld" (Deecke et al., 1985). There is also a large movement-related response 90–130 ms after EMG onset designated the "movement-evoked field" (MEF). Cheyne et al. (1989) found systematic variations in the location of the equivalent current dipole generators for both the RF and the MEF when face, hand, index finger, and thumb movements were studied. Their data are consistent with the motor homunculus, although explicit MRI correlation is lacking.

Advances in Evoked Magnetic Responses

Studies of evoked magnetic fields hold promise for increasing our knowledge of cerebral sensorimotor, and possibly cognitive, processes. Localization of sources by MEG is probably less affected by volume currents than localization based on EEG. In some cases, underlying ambiguity of sources and their orientation may be better evinced with MEG because of its sensitivity to tangentially (as opposed to radially) oriented current sources. A problem with many of the studies is that responses are sensitive to changing cognitive states. The present direction of these studies includes the measurement of responses to more complex stimuli (i.e., oddball paradigms), the comparison of evoked response information with that obtained by imaging modalities, and the incorporation of greater sophistication in source modeling, as in modeling spatiotemporal sources, as described by Scherg et al. (1989) and Baumgartner et al. (1989). Future directions that seem appropriate include (1) the development of larger, multiarray systems so that responses can be recorded simultaneously over the entire head, thus avoiding the unwanted variation in cognitive state; (2) the development of algorithms to utilize EEG and MEG information for source localization; and (3) the use of more realistic source and volume conductor models (i.e., based on MRI data). The clinical utility of these advances remains to be determined, but they clearly are of importance in our understanding of the mechanisms of brain function.

PRESURGICAL EVALUATION OF EPILEPTIC PATIENTS

Although MEG has been used to study a rather large number of epileptic patients worldwide, it is not an established diagnostic tool. Furthermore, MEG

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

currently does not influence the decision-making for the surgical treatment of epilepsy. When larger systems with faster and more realistic methods of analysis become available, the value of MEG may change significantly in this respect.

VALIDATION OF MEG

The point localization of sources predicted by MEG is unphysiological, so that the results cannot be compared with the widespread discharging region identified on ECoG recording. Invasive techniques such as depth or subdural electrode recordings are often used to validate the MEG and EEG findings (Sato et al., 1989), but spikes recorded with these methods are different from the scalp-recorded spikes in terms of amplitude, waveform, and distribution, and they cannot be compared on a one-to-one basis. Many small, independent spikes detected by depth or subdural recording may not be detected with the scalp-recorded EEG or MEG unless a wide area of the cortex fires synchronously.

An alternative way of comparing these three methods is to compare the areas explored by each method. ECoG delineates regions of epileptiform discharges on the cortical surface (Ajmone-Marsan, 1986; Rose et al., 1987*a*), whereas EEG and MEG may not easily provide regional information. Assuming that each spike represents a slightly different discharging region, analysis of many spikes will lead to an aggregate of many points, which in turn leads to a region (Sato et al., 1990). In this way, the discharging region identified with ECoG can be compared with that predicted by MEG.

FUTURE PERSPECTIVES IN MEG RESEARCH

Although the MEG localization of sources, especially deep beneath the skull surface, is not as affected by volume currents as a similar localization based on EEG mapping, MEG may have some difficulty in detecting deeper sources because of the quick decay of magnetic fields. MEG studies need to be compared with detailed EEG studies to assess the utility of each. Validation of each of the techniques must be performed in larger groups of patients treated surgically or studied with invasive techniques to assess the true utility of MEG.

A severe limitation of the currently available MEG systems in their inability to record events over the entire head simultaneously. This deficiency may introduce additional error, because nonsimultaneous events must be combined for analysis.

Without doubt, MEG is an excellent experimental tool but has yet to be proved a reliable clinical diagnostic methodology. Simultaneous EEG and MEG recording is always done to ensure that comparable signals at different locations are being compared; however, this also precludes the optimal use of MEG as an independent measurement. MEG technology is evolving, however, and a recent report from Finland (Tiihonen et al., 1990) describes a magnetometer that has seven sensors and a scanning area 93 mm in diameter. Furthermore, two 37-channel systems are now commercially available, and many other research institutes and commercial firms have expressed interest in producing even larger units to scan the entire head at once. It will be several more years until the routine clinical utility of MEG is established. Meanwhile, MEG and EEG will continue to complement each other, and both together will provide more information than either one alone.

Acknowledgment: The authors thank B. J. Hessie for editorial assistance.

REFERENCES

Aine C, George J, Medvick P, Supek S, Flynn E, Bodis-Wollner I. Identification of multiple sources in transient visual evoked neuromagnetic responses. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989:193–6.

Ajmone-Marsan C. Depth electrography and electrocorticography. In: Aminoff MJ, ed. *Electrodiagnosis in clinical neurology*, 2nd ed. New York: Churchill Livingstone, 1986: 205–30.

Antervo A, Hari R, Katila T, Poutanen T, Seppanen M, Tuomisto T. Cerebral magnetic fields preceding self-paced flexions of the foot. *Acta Neurol Scand* 1983;68:213–7.

Barth DS, Baumgartner C, Sutherling WW. Neuromagnetic field modeling of multiple brain regions producing interictal spikes in human epilepsy. *Electroencephalogr Clin Neurophysiol* 1989;73:389–402.

Barth DS, Sutherling W, Broffman J, Beatty J. Magnetic localization of a dipolar current source implanted in a sphere and a human cranium. *Electroencephalogr Clin Neurophysiol* 1986; 63:260–73.

Barth DS, Sutherling W, Engel J Jr, Beatty J. Neuromagnetic localization of epileptiform spike activity in the human brain. *Science* 1982;218:891–4.

Barth DS, Sutherling W, Engel J Jr, Beatty J. Neuromagnetic evidence of spatially distributed sources underlying epileptiform spikes in the human brain. *Science* 1984*a*;223:293–6.

Barth DS, Sutherling W, Engel J Jr, Beatty J. Neuromagnetic localization of single and multiple sources underlying epileptiform spikes in the human brain. In: Porter RJ, Mattson RH, Ward AA Jr, Dam H, eds. *XVth Epilepsy International Symposium.* New York: Raven Press, 1984*b*:379–84.

Baumgartner C, Barth DS, Sutherling WW. Spatiotemporal modeling of somatosensory evoked magnetic fields. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989:161–4.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Biomagnetic Technologies, Inc. *Marketing specification for 700 Series Magnetometer.* 1989.
Carelli P, Foglietti V, Modena I, Romani GL. Magnetic study of the spontaneous brain activity of normal subjects. *Il Nuovo Cimento* 1983;2D:538–46.
Chapman RM. The relative covariance method. In: Erne SN, Romani GL, eds. *Advances in biomagnetism. Functional localization: a challenge for biomagnetism.* Singapore: World Scientific, 1989:351–63.
Cheyne D, Kristeva R, Lang W, Lindinger G, Deecke L. Neuromagnetic localisation of sensorimotor cortex sources associated with voluntary movements in humans. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989:177–80.
Clarke CJS, Ioannides AA, Bolton JPR. Localised and distributed source solutions for the biomagnetic inverse problem I. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989:587–90.
Clarke J, Koch RH. The impact of high-temperature superconductivity on SQUID magnetometers. *Science* 1988:242:217–23.
Cohen D. Magnetoencephalography: evidence of magnetic fields produced by alpha rhythm currents. *Science* 1968;161:784–6.
Cohen D. Magnetic measurement and display of current generators in the brain. Part I: The 2-D detector. Presented at XIIth International Conference on Medical and Biological Engineering, Jerusalem, Israel, August 19–24, 1979.
Cohen D, Cuffin BN. Magnetic measurement and display of current generators in the brain. Part II: Polarization of the alpha rhythm. Presented at XIIth International Conference on Medical and Biological Engineering, Jerusalem, Israel, August 19–24, 1979.
Cohen D, Cuffin BN. Demonstration of useful differences between the magnetoencephalogram and electroencephalogram. *Electroencephalogr Clin Neurophysiol* 1983;56:38–51.
Cohen D, Hosaka H. Magnetic field produced by a current dipole. *J Electrocardiol* 1976;9:409–17.
Cuffin BN, Cohen D. Magnetic fields of a dipole in special volume conductor shapes. *IEEE Trans Biomed Eng* 1977*a*;24:372–81.
Cuffin BN, Cohen D. Magnetic fields produced by models of biological current sources. *J Appl Phys* 1977*b*;48:3971–80.
Deecke L, Boschert J, Brickett P, Weinberg H. Magnetoencephalographic evidence for possible supplementary motor area participation in human voluntary movement. In: Weinberg H, Stroink G, Katila T, eds. *Biomagnetism: applications and theory.* New York: Pergamon Press, 1985:369–72.
Ducla-Soares E, Rose DF, Sato S. A realistic head model for magnetoencephalography and the localization of epileptiform discharges. Presented at 18th Epilepsy International Symposium, New Delhi, India, October 14–18, 1989.
Elberling C, Bak C, Kofoed B, Lebech J, Saermark K. Auditory magnetic fields from the human cerebral cortex: location and strength of an equivalent current dipole. *Acta Neurol Scand* 1982;65:553–69.
Erne SN. SQUID sensors. In: Williamson SJ, Romani G, Kaufman L, Modena I, eds. *Biomagnetism: an interdisciplinary approach.* New York: Plenum Press, 1983:69–84.
Erne SN, Hoke M, Lutkenhoner B, Pantev C, Scheer H. Brainstem auditory evoked magnetic fields. In: Atsumi K, Kotani M, Ueno S, Katila T, Williamson SJ, eds. *Biomagnetism '87.* Tokyo: Tokyo Denki University Press, 1988:158–61.
Erne SN, Romani GL. Detection of magnetic evoked fields. In: Grandori F, Hoke M, Romani GL, eds. *Auditory evoked magnetic fields and electric potentials.* Basel: Karger, 1990:207–21. (Advances in audiology; vol 6.)
Freedman AP. Medical significance of the magnetic activities of the human body. In: Erne SN, Hahlbohm H-D, Lubbig H, eds. *Biomagnetism.* New York: Walter de Gruyter, 1981:153–76.
George JS, Aine CJ, Medvick PA, Flynn ER. Spatial/temporal resolution of multiple sources: paths of activation in human visual cortex. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989:197–200.
Gudden F, Hoening E, Reichenberger H, et al. A multi-channel system for use in biomagnetic diagnosis. *Electromedia* 1989;57:1–7.
Hari R. The neuromagnetic technique in the study of the human auditory cortex. In: Grandorfi F, Hoke M, Romani GL, eds. *Auditory evoked magnetic fields and electric potentials.* Basel: Karger, 1989*a*:222–82. (Advances in audiology; vol 6.)
Hari R. Activation of the human auditory cortex by various sound sequences: neuromagnetic studies. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989*b*:87–91.
Hari R, Aittoniemi K, Jarvinen ML, Katila T, Varpula T. Auditory evoked transient and sustained magnetic fields of the human brain. *Exp Brain Res* 1980;40:237–40.
Hari R, Ilmoniemi RJ. Cerebral magnetic fields. *CRC Crit Rev Biomed Eng* 1986;14:93–126.
Hari R, Kaukoranta E. Neuromagnetic studies of somatosensory system: principles and examples. *Prog Neurobiol* 1985;24:233–56.
Hari R, Lounasmaa OV. Recording and interpretation of cerebral magnetic fields. *Science* 1989;244:432–6.
Hughes JR, Cohen J, Mayman CI, Scholl ML, Hendrix DE. Relationship of the magnetoencephalogram to abnormal activity in the electroencephalogram. *J Neurol* 1977;217:79–93.
Hughes JR, Hendrix DE, Cohen J, et al. Relationship of the magnetoencephalogram to the electroencephalogram: normal wake and sleep activity. *Electroencephalogr Clin Neurophysiol* 1976;40:261–78.
Huttunen J. Magnetic cortical responses evoked by tactile stimulation of the middle finger in man. *Pflugers Arch* 1986;407:129–33.
Ilmoniemi RJ, Hamalainen MS, Knuutila J. The forward and inverse problems in the spherical model. In: Weinberg H, Stroink G, Katila T, eds. *Biomagnetism: applications and theory.* New York: Pergamon Press, 1985:278–82.
Ioannides AA, Bolton JPR, Hasson R, Clarke CJS. Localised and distributed source solutions for the biomagnetic inverse problem II. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989:591–4.
Kado H, Kashiwaya S, Higuchi M, Miura H. Direct approach to an inverse problem: a trial to describe signal sources by current elements distribution. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989:579–82.
Kaufman L, Williamson SJ. Magnetic location of cortical activity. *Ann NY Acad Sci* 1982;388:197–213.
Kouijzer WJJ, Stok CJ, Reits D, Peters MJ, Dunajski Z, Lopes da Silva FH. Neuromagnetic fields evoked by a patterned on-offset stimulus. *IEEE Trans Biomed Eng* MBE-32 1985;6:455–8.
Makela JP, Hari R, Leinonen L. Magnetic responses of the human auditory cortex to noise/square wave transitions. *Electroencephalogr Clin Neurophysiol* 1988;69:423–30.
Meijs JWH, Boom HBK, Peters MJ, Oosterom A. Application of the Richardson extrapolation in simulation studies of EEGs. *Med Biol Eng Comput* 1987;25:222–6.
Meijs JW, Peters MJ. The EEG and MEG: using a model of eccentric spheres to describe the head. *IEEE Trans Biomed Eng* 1987;34:913–20.
Mizutani Y, Okano K, Kuriki S. Measurement of neuromagnetic fields preceding triggered and self-paced finger movements.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989: 181-4.

Modena I, Ricci GB, Barbanera S, Leoni R, Romani GL, Carelli P. Biomagnetic measurements of spontaneous brain activity in epileptic patients. *Electroencephalogr Clin Neurophysiol* 1982; 54:622-8.

Nakasato N, Kado H, Nakanishi M, Koyanagi M, Kasai N, Niizuma H. Magnetic detection of sleep spindles in normal subjects. *Electroencephalogr Clin Neurophysiol* 1990:76:123-30.

Nunez P. *Electric fields of the brain.* New York: Oxford, 1981.

Nunez P. The brain's magnetic field: some effects of multiple sources on localization methods. *Electroencephalogr Clin Neurophysiol* 1986;63:75-82.

Pantev C, Hoke M, Lehnertz K, Lutkenhoner B, Anogianakis G, Wittkowski W. Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *Electroencephalogr Clin Neurophysiol* 1988;69:160-70.

Pantev C, Hoke M, Lehnertz K, Lutkenhoner B, Fahrendorf G, Stober U. Identification of sources of brain neuronal activity with high spatiotemporal resolution through combination of neuromagnetic source localization (NMSL) and magnetic resonance imaging (MRI). *Electroencephalogr Clin Neurophysiol* 1990;75:173-84.

Papanicolaou AC, Baumann S, Rogers RL, Saydjari C, Amparo EG, Eisenberg HM. Localization of auditory response sources using magnetoencephalography and magnetic resonance imaging. *Arch Neurol* 1990;47:33-7.

Reite M, Edrich J, Zimmerman JT, Zimmerman JE. Human magnetic auditory evoked fields. *Electroencephalogr Clin Neurophysiol* 1978;45:114-7.

Reite M, Teale P, Zimmerman J, Davis K, Whalen J. Source location of a 50 msec latency auditory evoked field component. *Electroencephalogr Clin Neurophysiol* 1988;70:490-8.

Reite M, Zimmerman JE, Edrich J, Zimmerman J. The human magnetoencephalogram: some EEG and related correlations. *Electroencephalogr Clin Neurophysiol* 1976:40:59-66.

Reite M, Zimmerman JT, Zimmerman JE. Magnetic auditory evoked fields: interhemispheric asymmetry. *Electroencephalogr Clin Neurophysiol* 1981;51:388-92.

Ricci GB. Clinical magnetoencephalography. *Il Nuovo Cimento* 1983;2D:517-37.

Ricci GB, Leoni R, Romani GL, Campitelli F, Buoriomo S, Modena I. 3-D neuromagnetic localization of sources of interictal activity in cases of focal epilepsy. In: Weinberg H, Stroink G, Katila T, eds. *Biomagnetism: applications and theory.* New York: Pergamon Press, 1985:304-10.

Ricci GB, Modena I, Barbanera S, et al. Tridimensional biomagnetic localization of epileptic foci. *Acta Neurochir Suppl* 1984; 33:85-9.

Ricci GB, Romani GL, Salustri C, et al. Study of focal epilepsy by multichannel neuromagnetic measurements. *Electroencephalogr Clin Neurophysiol* 1987;66:358-68.

Rogers RL, Papanicolaou AC, Baumann S, Saydjari C, Eisenberg HM. Nonstationary dynamics of sequential magnetic dipole source changes associated with N100 auditory evoked responses. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989:105-8.

Rogers RL, Papanicolaou AC, Baumann SB, Saydjari C, Eisenberg HM. Neuromagnetic evidence of a dynamic excitation pattern generating the N100 auditory response. *Electroencephalogr Clin Neurophysiol* 1990:77:237-40.

Romani GL, Leoni R. Localization of cerebral sources by neuromagnetic measurements. In: Weinberg H, Stroink G, Katila T, eds. *Biomagnetism: applications and theory.* New York: Pergamon Press, 1985:205-20.

Romani GL, Williamson SJ, Kaufman L. Tonotopic organization of the human auditory cortex. *Science* 1982*a*;216:1339-40.

Romani GL, Williamson SJ, Kaufman L, Brenner D. Characterization of the human auditory cortex by the neuromagnetic method. *Exp Brain Res* 1982*b*;47:381-93.

Rose DF, Ducla-Soares E, Sato S. Improved accuracy of MEG localization in the temporal region with inclusion of volume effects. *Brain Topography* 1989*a*;1:175-81.

Rose DF, Ducla-Soares E, Sato S. Comparison of EEG and multichannel magnetoencephalography in a patient with simple partial seizure. *Epilepsia* 1989*b*;30:665.

Rose DF, Sato S, Smith PD, et al. Localization of magnetic interictal discharges in temporal lobe epilepsy. *Ann Neurol* 1987*a*; 22:348-54.

Rose DF, Sato S, Smith PD, White J. Modelling the temporal region in patients with temporal lobe epilepsy. *Phys Med Biol* 1987*b*;32:69-63.

Rose DF, Smith PD, Sato S. Magnetoencephalography and epilepsy research. *Science* 1987*c*;238:329-35.

Rossini PM, Cilli M, Narici L, et al. Short-latency somatosensory evoked responses to median nerve stimulation: differences between electric and magnetic scalp recordings. In: Atsumi K, Kotani M, Ueno S, Katila T, Williamson SJ, eds. *Biomagnetism '87.* Tokyo: Tokyo Denki University Press, 1988:178-81.

Salustri C, Chapman RM. A simple method for 3-dimensional localization of epileptic activity recorded by simultaneous EEG and MEG. *Electroencephalogr Clin Neurophysiol* 1989; 73:473-8.

Sams M, Hamalainen M, Antervo A, Kaukoranta E, Reinkainen K, Hari R. Cerebral neuromagnetic responses evoked by short auditory stimuli. *Electroencephalogr Clin Neurophysiol* 1985;61:254-66.

Sarvas J. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Phys Med Biol* 1987:32: 11-22.

Sato S, Buchanan S, Balish M, Robinson S, Connaughton P, Kufta C. Single magnetic spike mapping with a 37-channel magnetometer. *Neurology* 1990;40(Suppl 1):256.

Sato S, Rose DF, Kufta CV, Theodore W. Validation MEG findings in clinical epilepsy research. In: Erne S, Romani GL, eds. *Advances in biomagnetism. Functional localization: a challenge for biomagnetism.* Singapore: World Scientific, 1989: 315-27.

Sato S, Rose D, Porter R. Single spike mapping. In: Weinberg H, Stroink G, Katila T, eds. *Biomagnetism: applications and theory.* New York: Pergamon Press, 1985:261-3.

Scherg M, Hari R, Hamalainen M. Frequency-specific sources of the auditory N19-P30-P50 response detected by a multiple source analysis of evoked magnetic fields and potentials. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989:97-100.

Stok CJ. The inverse problem in EEG and MEG with application to visual evoked responses. Leiden: Krisp Repro Meppel, 1986:67-81.

Stok CJ, Meijs JWH, Peters MJ. Inverse solutions based on MEG and EEG applied to volume conductor analysis. *Phys Med Biol* 1987:32:99-104.

Stok CJ, Meijs JWH, Peters MJ, Rutten WLC. Model evaluation using electroencephalography and magnetoencephalography. *Acta Otolaryngol (Stockh)* 1986;Suppl 432:5-10.

Suk J, Cappell J, Ribary U, Yamamoto T, Llinas RR. Magnetic localization of somatically evoked responses in the human brain. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989: 165-8.

Sutherling WW, Crandall PH, Cahan LD, et al. The magnetic field of epileptic spikes agrees with intracranial localizations in complex partial epilepsy. *Neurology* 1988*a*;38:778-86.

Sutherling WW, Crandall PH, Darcey TM, Becker DP, Levesque MF, Barth DS. The magnetic and electric fields agree with

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

intracranial localizations of somatosensory cortex. *Neurology* 1988*b*;38:1705-14.

Sutherling WW, Crandall PH, Engel J Jr, et al. The magnetic field of complex partial seizures agrees with intracranial localizations. *Ann Neurol* 1987;21:548-58.

Tepley N, Barkley GL, Moran JE, Simkins RT, Welch KMA. Observation of spreading cortical depression in migraine patients. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism*. New York: Plenum Press, 1989: 327-30.

Teyler TJ, Cuffin BN, Cohen D. The visual evoked magnetoencephalogram. *Life Sci* 1975;17:683-92.

Tiihonen J, Hari R, Hamalainen M. Early deflections of cerebral magnetic responses to median nerve stimulation. *Electroencephalogr Clin Neurophysiol* 1989;74:290-6.

Tiihonen J, Hari R, Kajola M, Nousiainen U, Vapalahti M. Localization of epileptic foci using a large-area magnetometer and functional brain anatomy. *Ann Neurol* 1990;27:283-90.

Vvedensky VL, Ilmoniemi RJ, Kajola MS. Study of the alpha rhythm with a 4-channel SQUID magnetometer. *Med Biol Eng Comput* 1986;23(Suppl), Part 2:11-2.

Wikswo JP Jr, Roth BJ. Magnetic determination of the spatial extent of a single cortical current source: a theoretical analysis. *Electroencephalogr Clin Neurophysiol* 1988;69:266-76.

Williamson SJ, Kaufman L. Analysis of neuromagnetic signals. In: Gevins A, Remond A, eds. *Handbook of electroencephalography and clinical neurophysiology*, revised series, vol 1. Amsterdam: Elsevier, 1987:405-48.

Williamson SJ, Kaufman L. Theory of neuroelectric and neuromagnetic fields. In: Grandori F, Hoke M, Romani GL, eds. *Auditory evoked magnetic fields and electric potentials*. Basel: Karger, 1990:1-39. (Advances in audiology; vol 6.)

Williamson SJ, Wang J-Z, Ilmoniemi RJ. Method for locating sources of human activity. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism*. New York: Plenum Press, 1989:257-60.

Wood CC, Cohen D, Cuffin BN, Yarita M, Allison T. Electrical sources in human somatosensory cortex: identified by combined magnetic and potential recordings. *Science* 1985;227: 1051-3.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

intracranial localizations of somatosensory cortex. *Neurology* 1988*b*;38:1705-14.

Sutherling WW, Crandall PH, Engel J Jr, et al. The magnetic field of complex partial seizures agrees with intracranial localizations. *Ann Neurol* 1987;21:548-58.

Tepley N, Barkley GL, Moran JE, Simkins RT, Welch KMA. Observation of spreading cortical depression in migraine patients. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989: 327-30.

Teyler TJ, Cuffin BN, Cohen D. The visual evoked magnetoencephalogram. *Life Sci* 1975;17:683-92.

Tiihonen J, Hari R, Hamalainen M. Early deflections of cerebral magnetic responses to median nerve stimulation. *Electroencephalogr Clin Neurophysiol* 1989;74:290-6.

Tiihonen J, Hari R, Kajola M, Nousiainen U, Vapalahti M. Localization of epileptic foci using a large-area magnetometer and functional brain anatomy. *Ann Neurol* 1990;27:283-90.

Vvedensky VL, Ilmoniemi RJ, Kajola MS. Study of the alpha rhythm with a 4-channel SQUID magnetometer. *Med Biol Eng Comput* 1986;23(Suppl), Part 2:11-2.

Wikswo JP Jr, Roth BJ. Magnetic determination of the spatial extent of a single cortical current source: a theoretical analysis. *Electroencephalogr Clin Neurophysiol* 1988;69:266-76.

Williamson SJ, Kaufman L. Analysis of neuromagnetic signals. In: Gevins A, Remond A, eds. *Handbook of electroencephalography and clinical neurophysiology*, revised series, vol 1. Amsterdam: Elsevier, 1987:405-48.

Williamson SJ, Kaufman L. Theory of neuroelectric and neuromagnetic fields. In: Grandori F, Hoke M, Romani GL, eds. *Auditory evoked magnetic fields and electric potentials.* Basel: Karger, 1990:1-39. (Advances in audiology; vol 6.)

Williamson SJ, Wang J-Z, Ilmoniemi RJ. Method for locating sources of human activity. In: Williamson SJ, Hoke M, Stroink G, Kotani M, eds. *Advances in biomagnetism.* New York: Plenum Press, 1989:257-60.

Wood CC, Cohen D, Cuffin BN, Yarita M, Allison T. Electrical sources in human somatosensory cortex: identified by combined magnetic and potential recordings. *Science* 1985;227: 1051-3.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Electroencephalography and clinical Neurophysiology, 1987, 67: 271–288
Elsevier Scientific Publishers Ireland, Ltd.

EEG 01898

EEG alpha map series: brain micro-states by space-oriented adaptive segmentation [1]

D. Lehmann, H. Ozaki [2] and I. Pal [3]

Department of Neurology, University Hospital, 8091 Zurich (Switzerland)

(Accepted for publication: 18 December, 1986)

Summary The spontaneous EEG, viewed as a series of momentary scalp field maps, shows stable map configurations (of periodically reversed polarity) for varying durations, and discontinuous changes of the configurations. For adaptive segmentation of map series into spatially stationary epochs, the maps at the times of maximal map relief are selected and spatially described by the two locations of maximal and minimal (extreme) potentials; a segment ends if over time an extreme leaves its pre-set spatial window. Over 6 subjects, the resting alpha EEG showed 210 msec mean segment duration; segments longer than 323 msec covered 50% of the total time; the most prominent segment class (1.5% of all classes) covered 20% of total time (prominence varied strongly over classes; not all possible classes occurred). Spectral power and phase of averages of adaptive and pre-determined segments demonstrated the adequacy of the strategy, and the homogeneity of adaptive segment classes by their reduced within-class variance. It is suggested that different segment classes manifest different brain functional states exerting different effects on information processing. The spatially stationary segments might be basic building blocks of brain information processing, possibly operationalizing consciousness time and offering a common phenomenology for spontaneous activity and event-related potentials. The functional significance of segments might be modes or steps of information processing or performance, tested, e.g., as reaction time.

Key words: Adaptive EEG segmentation; Spatial characteristics; Multichannel EEG segmentation; Alpha EEG momentary maps; Spectral analysis Nyquist diagrams

The functional state of the brain as manifested in the EEG determines the fate of information which is processed while the state exists, for example in wakefulness and sleep (see Koukkou and Lehmann 1983). The functional state is constrained by several factors (Koukkou et al. 1980); gross constraints are maturational stages (EEG: Katada et al. 1981), periodic circadian and ultradian processes (reaction time: Williams et al. 1966; brain electrical responses: Halász et al. 1985), and metabolic, drug, and disease conditions. The functional state is re-adjusted continuously within the constrained range in much shorter intervals for optimal processing of the information which is momentarily being treated (basically via the orienting response, see Ohman 1979; Rohrbaugh 1984), and the functional state varies as spontaneous short-term fluctuations (see 'fluctuations of attention' in Woodworth and Schlosberg 1954) in the realm of seconds or less. Various relationships between EEG data and short-term fluctuations of perception, vigilance and brain electrical responses have been shown using single-channel EEG data assessment (e.g., Lehmann et al. 1965; Keesey and Nichols 1967; Bohdanecky et al. 1984)

[1] Supported in part by the Swiss National Science Foundation, the Hartmann-Mueller Foundation and the EMDO Foundation, Zurich.
[2] Supported by Fellowships from the Swiss National Science Foundation, the Hoffmann LaRoche Research Foundation, Basel, and the Sleep and Dream Laboratory, Klinik am Zürichberg, Zurich. Home institution: Laboratory of Physiology for the Handicapped, Ibaraki University, Mito 310, Japan.
[3] Supported by a Fellowship from the Swiss National Science Foundation. Home institution: Computing Group, Semmelweis University, Budapest, Hungary.

Correspondence to: Dr. D. Lehmann, Neurology, University Hospital, 8091 Zurich (Switzerland).

0013-4649/87/$03.50 © 1987 Elsevier Scientific Publishers Ireland, Ltd.

Summary: p272 Good reference about extrema representing different sources
Rest of Article to much into Alpha but shows extrema

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

and adaptive segmentation of single channel EEGs (Gath et al. 1983, 1985).

The information in one EEG channel is a small fraction of the entire EEG scalp field data. Data from N electrodes can be combined into

$N * (N - 1)/2$

different EEG derivations (recording channels), i.e., 120 different wave shapes can be produced with 16 electrodes (Lehmann 1984; Lehmann et al. 1986). Adaptive time segmentation of the EEG into stationary epochs in the range of seconds has used time-oriented approaches and typically analyses data in one channel (Barlow et al. 1981; Lopes da Silva 1981; Gath et al. 1983). Different channels yield different segment boundaries. For more than one channel, a combination of the informations from the individually treated channels is proposed; up to 4 channels have been used (Jansen et al. 1982; Barlow 1985; Creutzfeldt et al. 1985).

Conceptualizing functional states as global conditions which involve the entire brain (Ashby 1960; Koukkou and Lehmann 1983; Wright et al. 1985), the scalp field ought to be considered as a whole. Accordingly, for the identification of different states, information from all locations ought to be used with equal weight without pre-selecting a sub-set of the available data. This can be achieved by using the spatial configurations of the series of momentary EEG field maps as unbiased information for adaptive time segmentation of global brain states. The underlying rationale is as follows: different spatial configurations of the scalp field maps must have been produced by the activity of different neural generator populations (Sidman et al. 1978; Grandori 1984), hence they are candidates for functionally different brain states. Identical field configurations, on the other hand, might reflect the activity of the same neural generators and therefore, of functionally similar states. Hence, the task is to determine epochs of variable length during which the configuration of the scalp field maps is stationary. Given the periodic nature of the spontaneous activity of neural populations with their repetitive polarity reversals, only the maps' spatial configuration, not their polarity, should be used as descriptor of the states.

Spatial field maps - different states

To study the functional states in detail, the present paper uses a method of adaptively segmenting EEG map series into time segments of variable length and stationary spatial characteristics. Since there is virtually no literature on the micro-phenomenology of spontaneous alpha EEG map series, the following section reviews some of their basic properties which are the basis of segmentation approaches.

Characteristics of alpha EEG map series. Fig. 1 shows an example of such a map series during two alpha cycles. The maps are basically simple (Lehmann 1971, 1975) and tend to be concentric around two extreme potential values, a maximal and a minimal, in some cases around 3 extrema. The configuration of momentary maps is not dependent on the chosen reference, only the voltage values of the contour lines or colours of the contour intervals depend on it. Aiming at spatial data reduction and feature extraction, a map's basic spatial characteristics can be given by the locations of the peak and trough potentials (the extrema locations) as shown in Fig. 2. This 2-point characterization of a map can be used with any number of electrodes and is analogous to describing a potential field by one equivalent model dipole. The characterization of a map by its extrema locations implies that all recording points are considered with equal weight, and that the extracted descriptors do not depend on the choice of the reference.

The accumulated extrema locations from a time series of successive alpha EEG maps exhibit a distinct spatial distribution (example Fig. 4B): the extrema occur frequently in restricted scalp areas at few electrodes, and infrequently at many other electrodes (see also Fig. 9 in Lehmann 1971 and Figs. 15 and 18 in Lehmann 1981). This map of the occurrence of field extrema over time is, as expected (Lehmann 1984; Lehmann et al. 1986), very similar to the map of spectral power of the EEG wave forms during the same analysis epoch if the average reference is used (Fig. 4C). The unequal spatial distribution of the extrema locations over time is the consequence of the tendency of extreme values, and thus of map configurations, to change quickly in a non-continuous, jump-like manner to different configurations (Lehmann

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

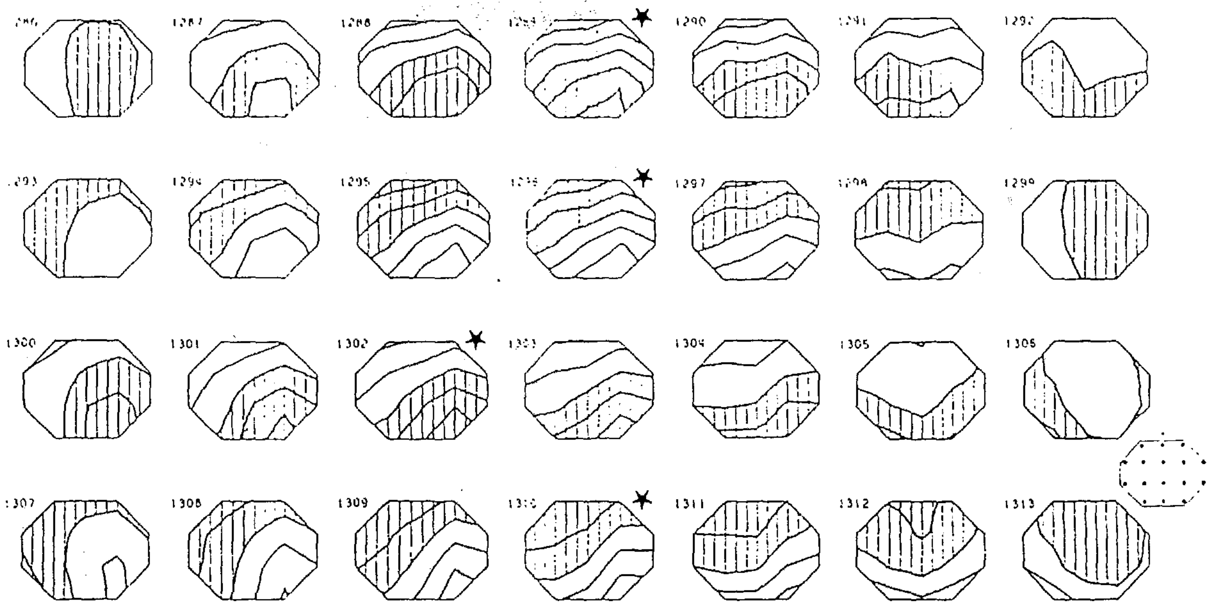


Fig. 1. Series of momentary isopotential contour maps of the scalp field distributions recorded from 16 electrodes during two cycles of alpha activity (epoch 'A' in Fig. 7) covering 211 msec. Maps at intervals of 7.8 msec (128 maps/sec), numbered consecutively. White positive, hatched negative relative to average reference; isopotential contour lines in steps of 10 μV; head seen from rear, left ear left, electrode array see schematic; most anterior electrode at Cz, most posterior 0.5 cm above inion at about equidistances. Asterisks mark the 4 maps of maximal relief (maximal global field power); see also Fig. 3. Note that polarity reverses for successive maps of maximal global field power, but that their spatial configuration remains stable. Fig. 2 (below) displays the major spatial characteristics of this map series.

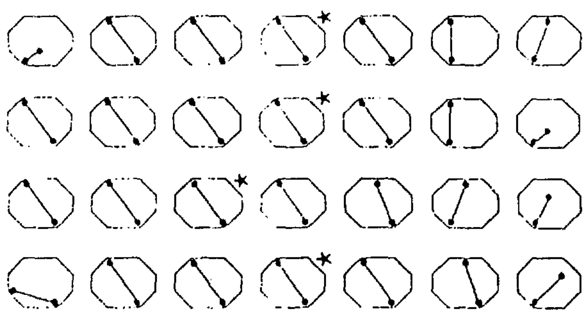


Fig. 2. Feature extraction of the major spatial characteristics of the map series in Fig. 1. The configuration of each map is characterized by the locations of the positive and negative extreme potential values which are indicated by dots and connected by a line. Maps at times of maximal global field power marked by asterisks. Note that the configurations of the maps remain stable for several successive maps around the times of maximal global field power (see also Fig. 3), and change stepwise during times of low field power.

1971, 1981, 1984). The momentary maps show no wave fronts and their major characteristics do not propagate or travel continuously over large field distances (Lehmann 1971, 1981).

The series of maps in Fig. 1 illustrates that twice during a spontaneous alpha cycle there are large differences between the potentials at the most positive and the most negative locations, as indicated by a larger number of isopotential lines; and that at two other times, the maps tend to be flat with few isopotential lines (Lehmann 1971). Using the values at all electrodes, the degree of relief or hilliness in a map can be assessed numerically by the 'hilliness index' (Lehmann 1971) or the directly related measure 'global field power' (Lehmann and Skrandies 1980); these are reference-independent.

Global field power (formula 1 in Appendix) of the data shown in Fig. 1 is plotted as a function of time in Fig. 3. For the alpha band, global field power typically shows maximal values about every

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

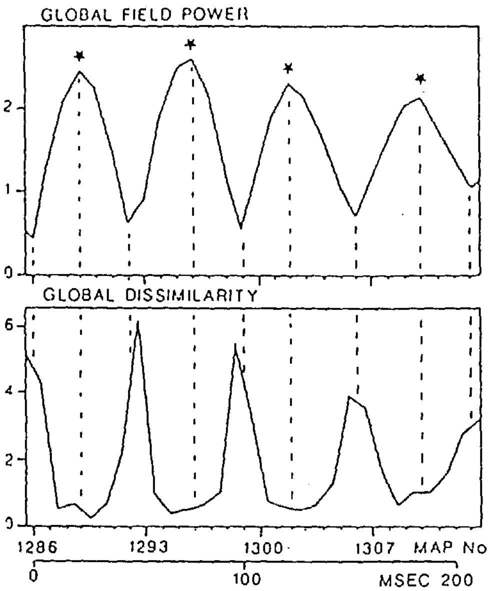


Fig. 3. Global field power per map and global dissimilarity of successive maps of the 211 msec map series of two cycles of alpha activity shown in Figs. 1 and 2. Times of maximal global field power marked by asterisks. Note longer periods of low dissimilarity, i.e., stability of the spatial configurations of the maps during times of high global field power, and shorter periods of changes of map configurations during low field power. Vertical scales in arbitrary units. Map numbering corresponds to Fig. 1.

50 msec (Fig. 3; see also Fig. 7, bottom). The time between two maxima of global field power corresponds to a half-wave of conventional voltage wave form records (see also Fig. 1), since there is one maximal and one minimal value of global field power for each half cycle of a rhythmic brain activity (Lehmann 1971). Global field power is a one-number statement for all electrodes, and the actual location of the map's peak or trough is unimportant for the detection of the 20/sec maxima of global field power during 10/sec alpha activity.

In the most basic case the maps at successive times of maximal field power during an alpha cycle are similar in configuration, but reversed in polarity (cf., the maps marked with asterisks in Fig. 1). It will be seen that in other cases there is a change of map configuration between two successive times of maximal field power.

The difference in spatial configuration between maps can be assessed by the one-number measure (formula 2 in Appendix) of 'global dissimilarity' (Lehmann and Skrandies 1980). Used on successive pairs of maps, global dissimilarity as a function of time indicates periods of stability and instability of map configurations.

Fig. 4A and B illustrate that the map configurations do not change continuously from map to map, but rather change stepwise, being stable for relatively long times. Figs. 1 and 3 illustrate that stable configurations are typically seen during times of high global field power, and that changes to new map configurations, which are manifest as high values of global dissimilarity, occur quickly, typically during times of low global field power (Lehmann 1971, 1981; Lehmann et al. 1986).

Accordingly, exclusive sampling of extrema locations (data reduction in space via feature extrac-

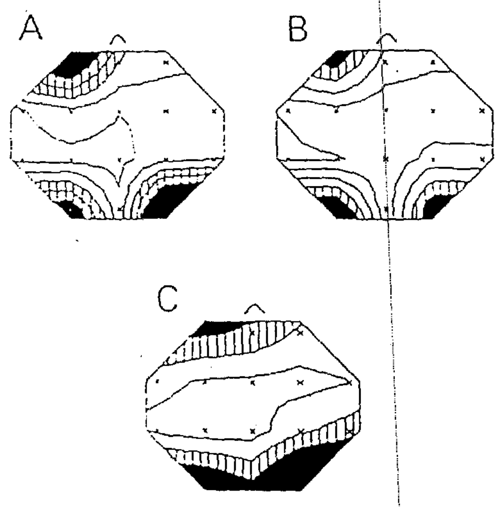


Fig. 4. Mapping of digitally bandpassed 8–12 Hz alpha activity during an epoch of 24 sec duration. A: map of the spatial distribution of the extrema locations (positive and negative in each map) accumulated from the 461 maps at the times of maximal global field power during the epoch. B: map of the spatial distribution of the extreme potential locations accumulated from all momentary maps (128/sec; 3072 maps) during the epoch. C: map of the spatial distribution of the spectral band power density of the EEG wave forms vs. the average reference during the epoch. High values black, medium values hatched, low values white. Contour lines connect interpolated equal values of occurrence frequency or spectral band power. Crosses are electrode locations. Note the similarity of the three maps.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

tion) only at the moments of maximal global field power (data reduction in time) is expected to be representative of the entire epoch under study. This is supported by the similarity of the maps obtained with extrema locations from all sampling times, with extrema locations from only the times of maximal global field power, and with values from all locations at all times, as shown in Fig. 4A,B and C (see also later Fig. 11). Exclusive sampling at times of maximal field power has the additional advantage that extrema determination in these maps with high relief will be less vulnerable to noise than in maps with low relief.

The example of Fig. 5 shows a series of momentary maps selected at successive times of maximal global field power during spontaneous activity. In each map, the locations of the most positive and most negative potentials are marked. Obviously, these locations remain in restricted areas for several reversals of field polarity. This illustrates how the maps at times of maximal global field power tend to spatially stable configurations over successive alpha half cycles; after a quick change to a new map configuration, the maps again tend to be stable for several half cycles (see also Lehmann 1971, 1984). The segmentation procedure

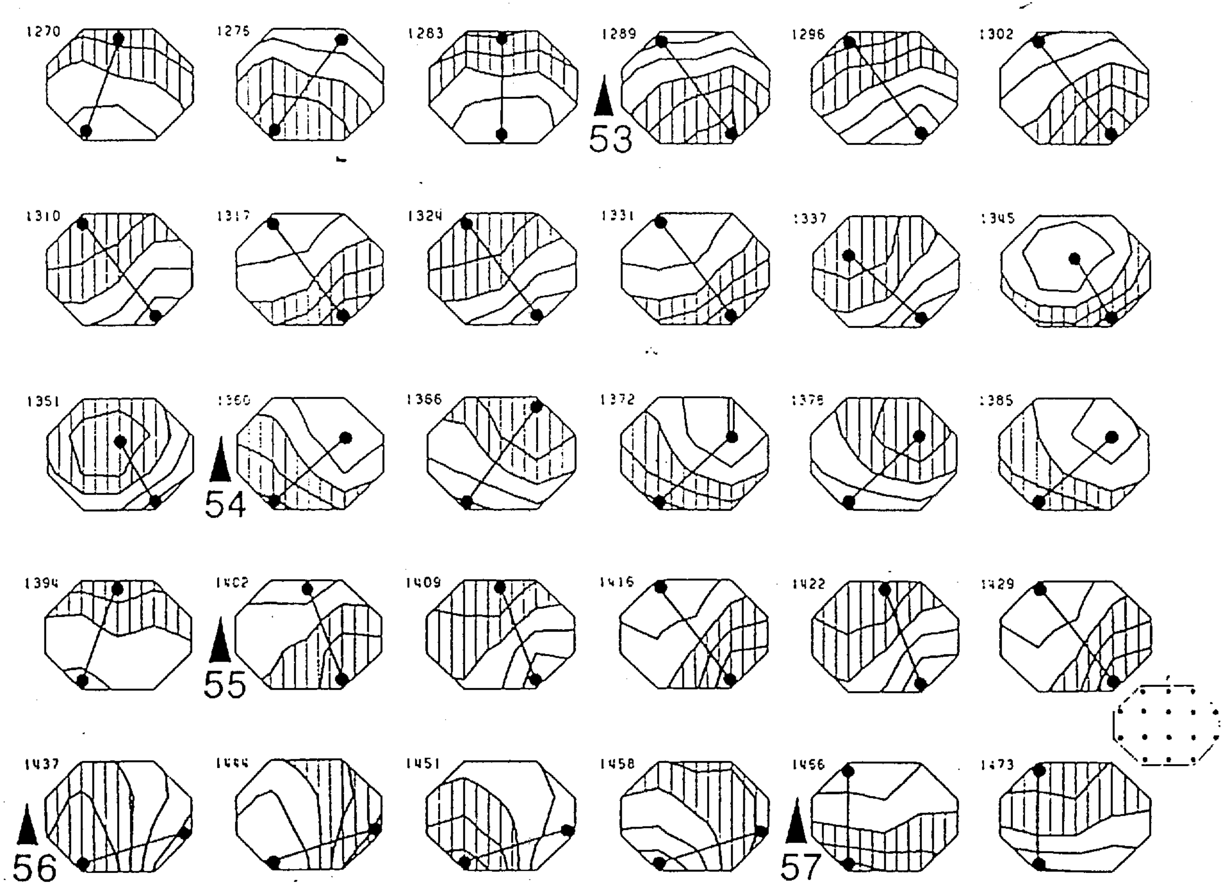


Fig. 5. Adaptive segmentation of a series of momentary scalp field maps covering 1585 msec of alpha activity (epoch 'B' of Fig. 7) recorded from 16 electrodes. The maps at successive times of maximal global field power are used. Average time between maps is 54.7 msec; numbers identify original sampling time points (128/sec); white positive, hatched negative relative to average reference; isopotential contour lines in steps of 10 μV. In each map, the locations of the maximal and minimal potentials are marked by dots and connected by a line. A 4-electrode square window is used for segmentation (see text). Segment terminations are marked by vertical arrows and the segment numbers correspond to those in Fig. 7. See Fig. 8 for maps of the segment classes.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

used in this paper is designed to identify these brief epochs of stability. It is hoped that the results clarify the micro-phenomenology of the EEG.

Methods

EEG data were obtained as 16-channel records from 6 normal, unselected, right-handed, male volunteers between 25 and 35 years of age. Grass gold cup electrodes were placed with Grass EC-2 cream at about equidistant spacing over an area between the vertex and 0.5 cm above the inion, using the array schematic shown in Fig. 6. Using one of the electrodes as common reference, the spontaneous EEG was recorded from the comfortably seated subjects in a sound-shielded chamber. After 5–8 min of adaptation to the recording surrounds, ADC started 30 sec after the request to close the eyes. After anti-alias filtering, the data were sampled for 2 min at 128 samples/sec, and digitally filtered off-line using a finite impulse-response linear phase filter (McClellan et al. 1973) with a bandpass of 8.6–11.6 Hz (3 dB points; 20 dB down at 7 and 13 Hz). The data were reformatted into series of momentary voltage maps for further analysis. For display and comparisons, the maps were recomputed vs. the average reference.

Segmentation

Global field power was computed for each map and examined over time. The maps at all time points of maximal global field power were selected and only these were used for segmentation. In these maps the locations of the positive and the negative field extreme potentials were determined. A segment was accepted to continue as long as the

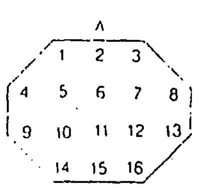


Fig. 6. Schematic of the array of electrodes (numbered). Head seen from rear, most anterior electrode at Cz, most posterior 0.5 cm above inion, about equal inter-electrode distances.

extrema of successive maps remained in the two spatial window areas which are defined by the locations of the extrema in the first maps of the segment. We used spatial windows consisting of a square array of 4 electrodes (but a segment might terminate without extrema having occurred at all 4 electrodes). The electrode where the extreme occurred in the first map of a segment is the initial location. If, in a subsequent map, an extreme occurs at a neighbor electrode in the transverse or sagittal direction, the window is defined in the respective direction, and if it occurs at a diagonal neighbor electrode, the window is defined in both directions. If an extreme occurs outside of the window, or farther away from the initial location than one electrode distance, the segment is terminated, and both windows are reset for the next segment. It is possible that the two windows of a segment overlap partially.

As an example, let one extreme in the first map occur at electrode 10 in Fig. 6. If in a subsequent map an extreme occurs at electrode 4 (or 6; or 15), the square window is completely defined, comprising electrodes 4, 5, 9 and 10 (or 5, 6, 10 and 11; or 10, 11, 14 and 15). If, after the initial location, the extreme occurs at electrodes 5 or 9 or 11 or 14, the window is defined in only one direction, and complete definition requires an additional occurrence location (if the extreme occurred at electrode 9 and 14, the window would consist of only 3 electrodes). As long as subsequent extremes occur within the set window, the segment continues, unless the other extreme left its window. An example of the application of these rules to a sequence of momentary maps at successive times of maximal global field power is shown in Fig. 5.

The class to which a segment belongs is defined by the two electrode locations (one in each spatial window) with the highest occurrence frequencies of extrema over all analyzed maps of the segment. Following the above-described window definition, 120 different segment classes are theoretically possible for our 16-electrode array. If a tie of occurrence frequencies existed between electrodes within a window area, i.e., if 2 or even 3 electrodes in the window had the same frequency of extreme occurrence during the segment's duration, then the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

higher sum of the global field power values of the momentary maps out of which the extrema were collected would decide the electrode which defined the segment class.

Results

A sample of adaptive segmentation of an alpha bandpassed recording of several seconds is illustrated in Fig. 7, where the segment boundaries are drawn into the EEG wave shapes (recomputed vs. the average reference) during the first 8 sec of the recording of one of the subjects. The unequal durations of the segments are evident. The segments which were found during the 4.7 sec epoch 'C' of the wave shapes of Fig. 7 are illustrated in Fig. 8. This figure gives for each segment an interpolated map of the accumulated extrema of all momentary maps at maximal field power times during the segment (these maps of extrema occurrence are of the same type as those in Fig. 4A and B). Fig. 8 shows that, in most segments, there is a very clear preponderance of the extreme occurrence at one of the 3 or 4 electrodes in the spatial

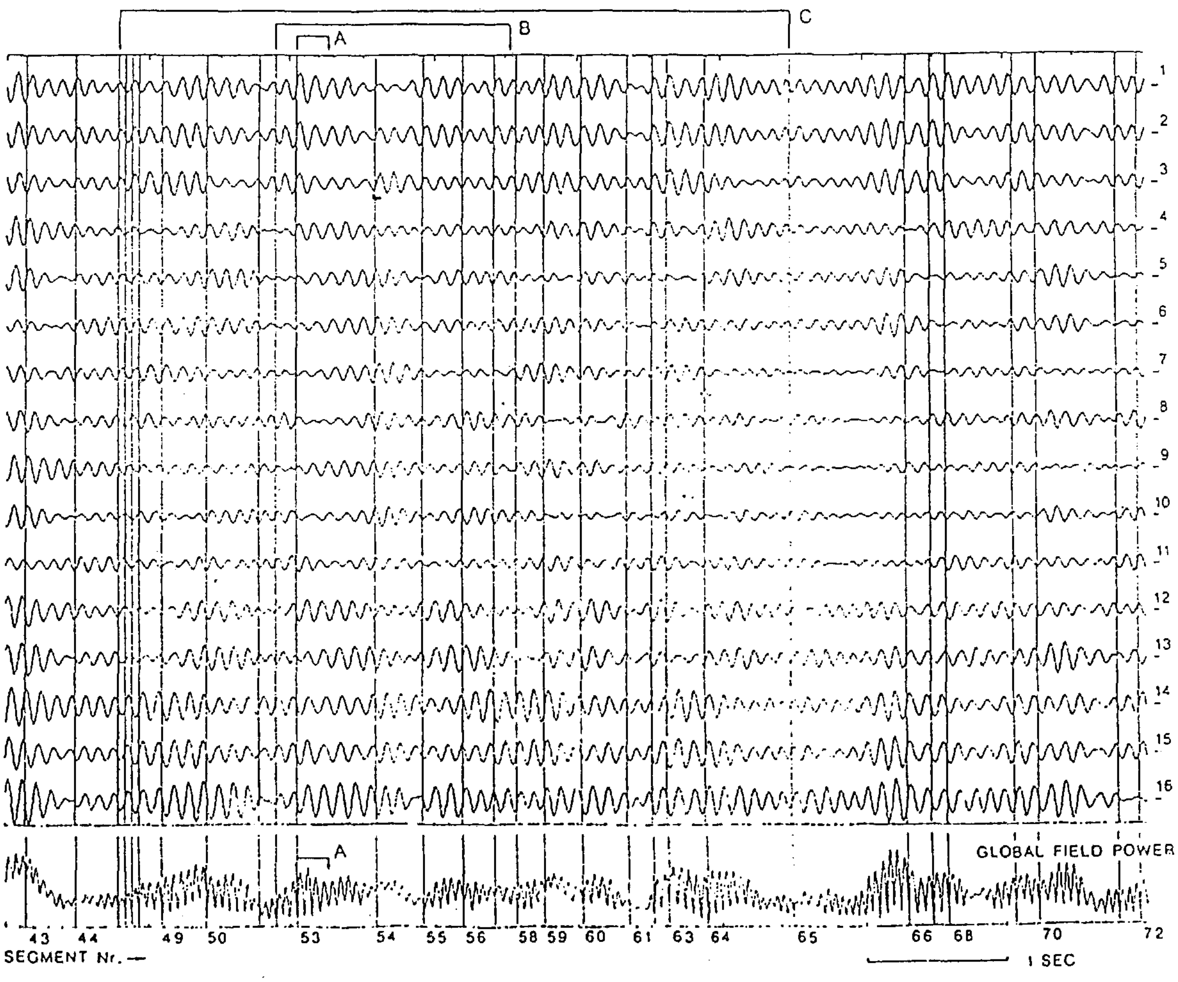


Fig. 7. Wave shapes vs. the average reference recorded during 8 sec in 16 channels (electrode array of Fig. 6) and digitally filtered to 8–12 Hz. Vertical distance between zero lines of successive channels is 90 μV. Bottom trace shows global field power. Vertical lines show the adaptively determined segment boundaries; segment numbers at bottom correspond to the numbers in Figs. 5 and 8. Epoch 'A' was used for Figs. 1, 2 and 3, 'B' for Fig. 5, and 'C' for Fig. 8.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

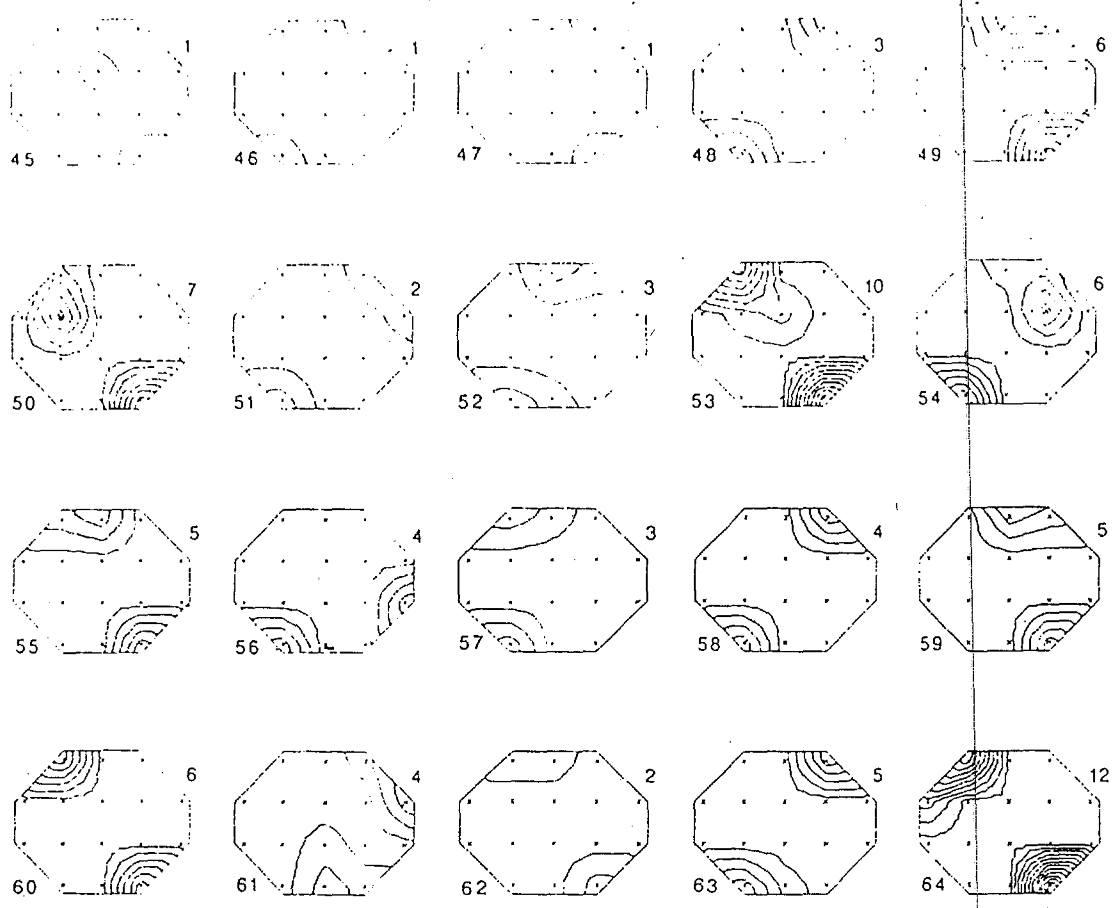


Fig. 8. Maps of occurrence of extreme potential values during the successive segments of the 4.7 sec epoch 'C' in Fig. 7. During each adaptively determined segment, the momentary maps at all times of maximal global field power were searched for the locations of the two extreme potentials; these locations were accumulated and linearly interpolated between electrodes to construct the present maps. (The number of iso-frequency-of-occurrence lines therefore is related to the number of times of maximal global field power.) The figure to the left below each map is the segment number used in Fig. 7. The figure to the right above each map is the number of times of maximal global field power during the segment; since maximal global field power occurs at about 50 msec intervals, the upper right figure multiplied by 50 indicates segment duration in msec.

window, even for relatively long segments. Only segments 49, 51 and 62 showed equal occurrence at 2 electrodes within a window, and had to be classed considering the summed field power values. There is no obvious rule of sequence for the different segment classes in Fig. 8. The successive segment classes are very different in spatial characteristics, e.g., the series of 10 'anterior left to posterior right' oriented map configurations of segment 53 in Fig. 8 is followed by the 6 'anterior right to posterior left' configurations of the next segment, which is followed by the 'anterior midline to posterior right' segment 55 and then by the 'posterior temporal right to posterior left' oriented segment 56.

Table I gives an overview of the major features of the segments for the individual subjects and the means for the sample population. On the average over subjects the segments' mean duration was 210 msec. The distribution of total time covered by segments of increasing duration was strongly skewed (Fig. 9); over subjects, about 70% of the total time was covered by segments of 210 msec duration or longer, 50% of total time by segments

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE I

Numerical characteristics of alpha band segments of 2 min analysis epochs in each of 6 subjects. Entries for individual subjects were rounded if greater than 3; means and standard deviations over subjects were computed before rounding.

	Subject no.						Mean (n = 6)	S.D.
	1	2	3	4	5	6		
(A) All segments during the analysed epoch of 2 min								
Duration of segments, mean (msec)	238	201	225	259	223	117	210	50
50% of total time was covered by segments longer than (msec)	389	310	336	387	360	155	323	88
25% of total time was covered by segments longer than (msec)	610	487	538	619	555	237	508	141
Number of segment classes in 2 min	52	65	60	45	53	85	60	14
Number of segments/min	252	298	268	232	270	516	305	104
Number of segments/class/min	5	5	4	5	5	6	5.0	0.6
(B) Segments of the class which covered maximal total time of all classes								
Duration, mean (msec)	471	329	337	403	377	152	345	108
% time covered	28	16	15	26	28	8	20.2	8.4
% of all segments	14	10	10	17	16	7	12.3	3.9
Number of segments/min	36	29	28	39	44	34	34.6	6.2
(C) Segments of the class covering second to maximal total time of all classes								
Duration, mean (msec)	307	243	265	258	331	192	266	49
% time covered	9	12	13	11	16	7	11.3	3.1
% of all segments	7	10	11	11	11	4	9.0	2.9
Number of segments/min	17	30	31	26	29	21	25.7	5.6
(D) Segments which lasted longer than 600 msec								
Duration, mean (msec)	852	834	831	848	933	609	817	109
% time covered	27	19	19	26	19	0.5	18.4	9.5
% of all segments	8	5	5	8	5	0.1	5.2	2.9
Number of segments/min	19	14	14	19	13	0.5	12.9	6.7
Number of segments/class/min	1.7	1.5	1.2	1.6	2.1	0.5	1.4	0.5
The segments belonged to how many classes	11	9	11	11	6	1	8.2	4.0
or, % of all classes	11	14	18	24	11	1.2	13.2	7.6
Of these segments over 600 msec, how many were over 1000 msec?/min	4	3	2	3	2	0	2.2	1.2
Longest segment (msec)	1910	1328	1638	1492	2656	609	1605	675

longer than 323 msec, and 25% by those longer than 508 msec. In all 6 subjects, only about 50% of the theoretically possible 120 different segment classes were actually found during the analyzed 2 min records. (With increased analysis time there is an expected, but levelling-off, increase in the number of classes, e.g., in subject no. 1: 28 classes in 30 sec, 47 in 60 sec and 52 in 120 sec.) In addition, the different segment classes showed very different frequencies of occurrence. On average, there were 5 segments/class/min, but the frequencies varied greatly: the segment class which covered most total time per subject ('most prominent' class, Table 1B) occurred about 35 times/min, and

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

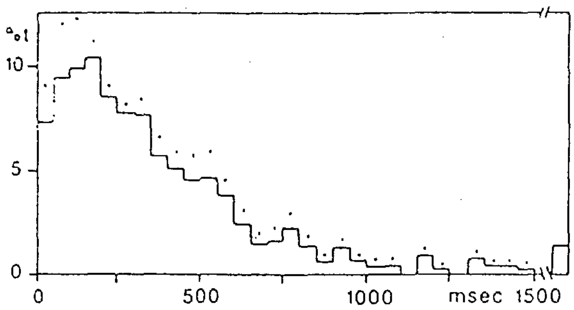


Fig. 9. Total % time (vertical) covered by segments of different durations (horizontal). Mean and S.E. over the 6 subjects.

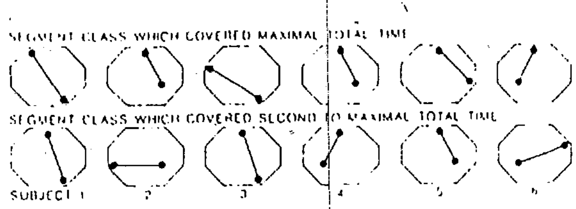


Fig. 10. Schematic of segment classes which covered most time (upper row) and second most time (lower row) in the record of each subject.

the class which covered the next most time (Table IC) occurred about 26 times/min. These two most prominent classes also had significantly increased mean segment durations over subjects (paired Wilcoxon P smaller than 0.025), longer by more than 50% and 25%, respectively, than the average segment duration. However, long segments over 600 msec, which constituted only 5% of all segments, belonged to many different classes (Table ID) and showed an average occurrence per class of only 1.5 segments/class/min.

The most prominent segment class on the average over subjects covered about 20% of the total time with only about 12% of all segments, and the two most prominent classes together covered over 30% of the time with 21% of all segments. The classes which covered most time (Fig. 10) in subjects 1 through 5 were quite similar, belonging to the 'anterior left (or midline) to posterior right' type, which reflects the frequent right hemisphere alpha dominance; only subject 6 and a most prominent class of the 'anterior midline to posterior left' type. Fig. 10 shows also that the classes which covered the second most total time differed more between subjects. In all but subject 4, the longest segment belonged to one of the two most prominent classes.

Our segment class definition makes it possible that successive segments might belong to the same class. This, however, occurred only 28 times (0.08%) among the 3666 segment boundaries found in the present entire material.

The segment boundaries in our procedure might depend on the starting point of the analysis, since the first occurrence of an extreme outside the initial location decides the location of the window. This problem is not serious since the segmentation becomes unambiguous after the first boundary which is caused by a change of an extreme location by more than one electrode, an event which occurs very frequently, in the present material in nearly 40% of the segments.

Validation

For feature extraction and segmentation, we applied a data reduction in time by selecting the maps at the times of maximal global field power, and in space by selecting the two extrema locations in each map as crucial features. This characterization of the segments by two extrema omits map characteristics of secondary strength. We examined the possible importance of the omitted information by comparing the two-extrema description of adaptively determined segments with power maps of the same segments computed from all originally available data. We also tested whether adaptive segmentation which is based on the reduced data set successfully leads to reduced variance over time and increased variance over space, i.e., to more stationary and more pronounced average map relief than pre-determined segmentation.

Fourier transforms of the complete data sets of adaptively determined segments were computed. Epoch length for transformation was 594 msec, covering 76 data points; a 1/10 cosine taper window preceded FFT, and the spectra were 3-point-smoothed (1/4, 1/2, 1/4). Power was computed vs. the average reference, i.e., using spatial DC rejection, a meaningful and physiologically interpretable way to examine local variance of

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

field potentials over time (Fig. 4B; see also Lehmann 1971, 1984; Walter et al. 1984; Bertrand et al. 1985; Lehmann et al. 1986). Fig. 11 shows maps of the 10.1 Hz spectral power of the wave shapes of the adaptively determined segments. These power maps of the complete data set are compared with our 2-location descriptions of the same data.

The power maps vs. the average reference of the individual adaptively determined segments are very similar within a given class and reflect their respective classes well (Fig. 11A and B), while those of the pre-determined fixed-length consecutive segments (Fig. 11C) differ considerably one from another. However, even though mean maps of spectral band power of wave shapes of longer data epochs (example Fig. 4C) almost always show 3 spatial peaks of power, very short, pre-determined segments such as in Fig. 11C sometimes have only two spatial peaks of power, probably because much of the short analysis epoch is covered by one class of adaptively determined segments. The example of averaged sine-cosine vector diagrams (Lehmann et al. 1986) in Fig. 12 illustrates that the variance of phase angles between different electrodes over segments is much smaller (size of circles around electrode entries) for adaptively determined than for pre-determined segments. The adaptively segmented data in these vector diagrams also show a more distinct configuration of the entries, with two electrodes clearly at the extreme positions (in Fig. 12A: electrodes 16 and 4, and Fig. 12B: electrodes 16 and 2) and thus showing maximal power of the wave shapes over time, whereas in the pre-determined segment averages, several electrodes are close to the extreme positions and show overlapping standard deviations, indicating a less distinct spatial distribution of spectral power computed over time.

When comparing adaptively determined segments (Table IIA) and predetermined segments (Table IIB and C), there are 3 expectations if one

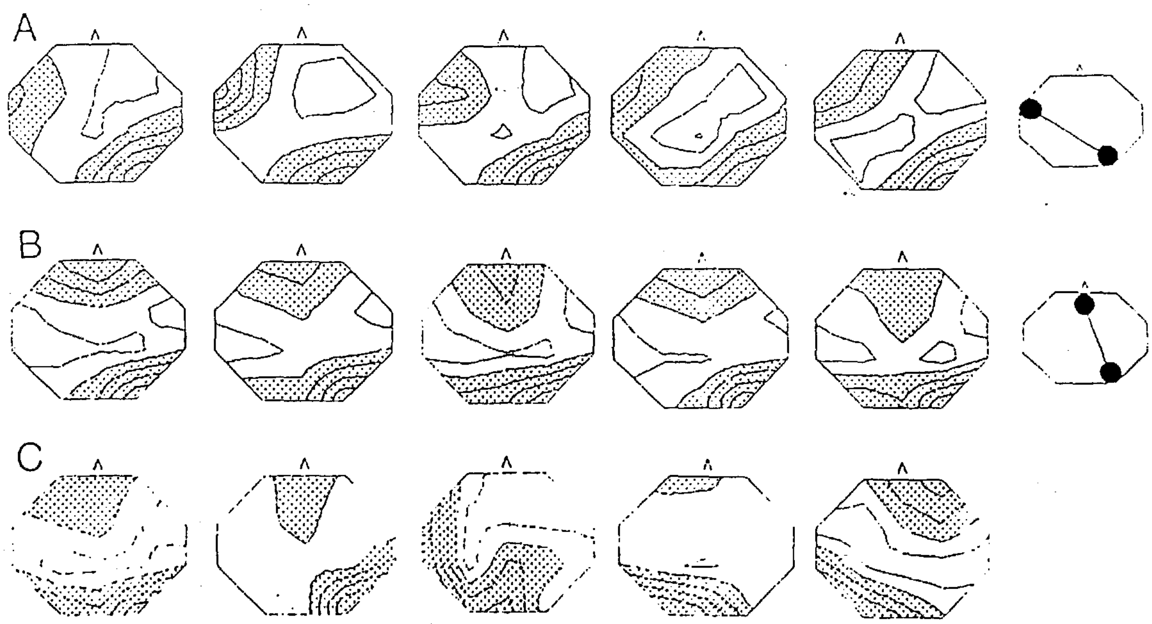


Fig. 11. Iso-power contour maps of the 10.1 Hz point of the Fourier-transformed wave shapes vs. the average reference of data segments of 594 msec duration, from one subject. A: power maps of 5 adaptively determined segments of class 'a' (schematic on the right). B: power maps of 5 adaptively determined segments of class 'b.' C: power maps of the first 5 successive pre-determined segments of the epoch used in Table IIC. High power stippled, low power white; iso-power contours in equal steps. Note consistency of spatial configuration of power maps of adaptively determined segments and their agreement with the schematic (segment class), and inconsistency between pre-determined segments.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

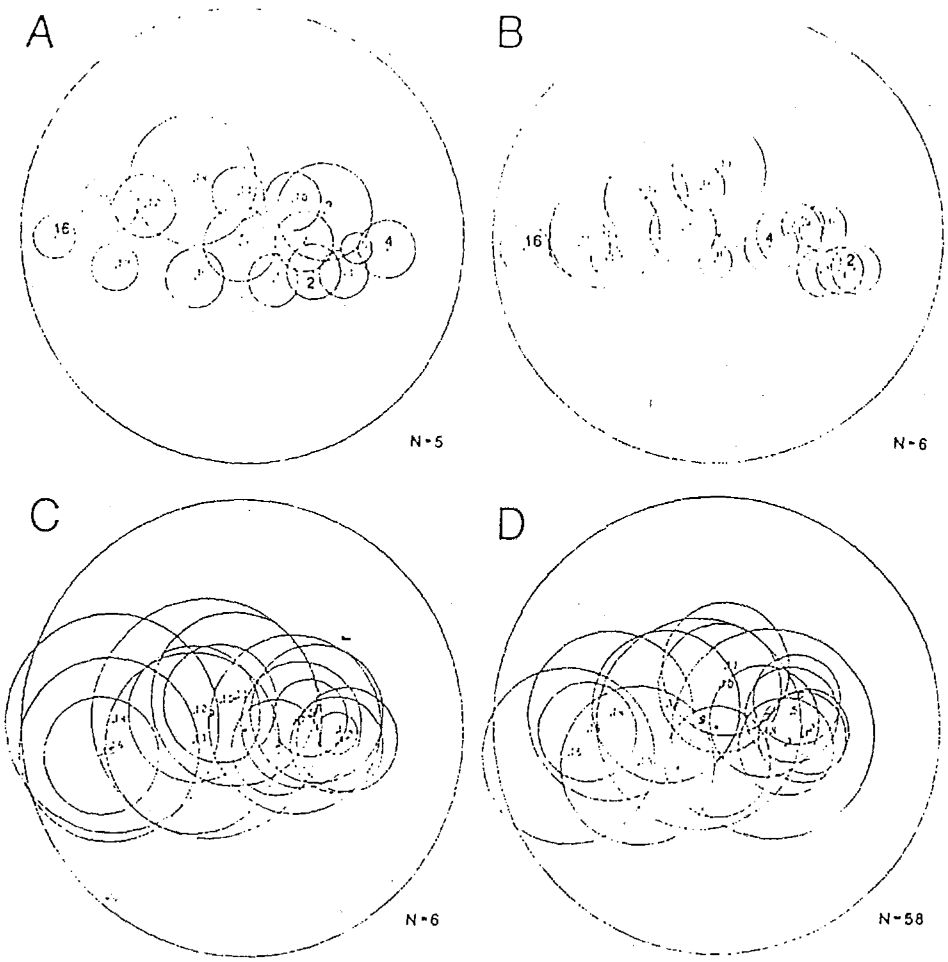


Fig. 12. Sine-cosine vector diagrams (Nyquist diagrams) of the 10.1 Hz point of 16-channel wave shapes, Fourier-transformed in segments of 594 msec duration and averaged over N epochs using the best fit rotation method (Lehmann et al. 1986). Cosine coefficient horizontal, sine coefficient vertical. Numbers of entries indicate electrode positions (see Fig. 6); heavy numbers mark electrodes with maximal power in A and B. Circles give the standard deviation of the entries. The center of the diagram is the average reference point, the mean of all entries. A: 5 adaptively determined segments of class 'a' were averaged. B: 6 adaptively determined segments of class 'b' averaged. C: 6 successive, pre-determined segments were averaged. D: average of the first 58 successive pre-determined segments of a continuous recording. Note larger standard deviations around average values of pre-determined segments than adaptively determined segments.

considers all sampling points in time and space: (1) Within a given class of adaptively determined segments, the variance of the configuration of maps of spectral band power of EEG wave shapes (Table II, column 5) should be minimal over segments (even though they occur non-consecutively), and it should be smaller than the variance of power map configurations over successive, pre-determined segments. (2) It follows from this that the relief of the mean power map over the power maps of several segments (Table II, column 3) should be more pronounced for adaptively determined segments of a given class than for pre-determined segments. (3) It also follows that for adaptively determined segments the mean value of relief of the power maps of all segments of a class (Table II, column 4) ideally should be similar to the relief value of the mean power map of the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

TABLE II

Characteristics of configurations of maps of the 10.1 Hz spectral power of the wave shapes of data epochs (segments) of 594 msec duration, from the 2 min recording of one subject. The maps were normalized for unity relief magnitude in order to focus on map configuration. (A) Mean power maps of all adaptively determined segments which belonged to the 4 classes of segments with three or more segments of 594 or more msec duration during the recording time. (B) Mean power maps of 4 groups (Ns chosen to match Ns in A) of segments which were randomly selected from the 58 consecutive, pre-determined segments in C. (C) Mean power map of the first 58 pre-determined segments of the recording. The computation of map relief and dissimilarity was done after formulae (1) and (2) in the text.

(1) Class/group of segments	(2) Segments (N)	(3) Relief (field power) of the mean map of the N maps	(4) Mean relief (field power) of the reliefs of the N maps	(5) Dissimilarity between the individual maps which contributed to the mean map
(A) Mean 10.1 Hz power maps of 4 different classes of adaptively determined segments				
1.	N = 5	0.259	0.262	0.108
2.	N = 6	0.249	0.262	0.105
3.	N = 3	0.232	0.236	0.048
4.	N = 3	0.282	0.299	0.087
Average	(N = 4)	0.256 ***	0.265 +++	0.087 ***
(S.D.)		(0.0209)	(0.0259)	(0.0276)
(B) Mean 10.1 Hz power maps of 4 groups (Ns matched to data in A) of pre-determined segments which were randomly selected from C				
1.	N = 5	0.165	0.250	0.206
2.	N = 6	0.196	0.272	0.190
3.	N = 3	0.187	0.229	0.139
4.	N = 3	0.168	0.258	0.171
Average	(N = 4)	0.179 ***	0.252 +++	0.177 ***
(S.D.)		(0.0149)	(0.0179)	(0.0288)
(C) Mean 10.1 Hz power map of 58 consecutive, pre-determined segments				
1.	N = 58	0.159	0.258 +++	0.201
(S.D.)			(0.0261)	
	(N = 1)	0.159 ***		0.201 ***
(S.D.)		(0)		(0)

*** The average relief and dissimilarity values of class mean maps of adaptively determined segments (A) are significantly different from the corresponding values of the randomly selected pre-determined segments (B) and from those of all 58 pre-determined segments (C), with all t test P levels smaller than 0.005 (non-pooled variances, $df = 6$ and $df = 3$, respectively).

+++ The difference between the relief of the mean map (column 3) and the mean relief of all N maps (column 4) is significantly smaller for the adaptively determined segments (A) than for the pre-determined segments (B) and (C); both t test P values (A vs. B and A vs. C) were smaller than 0.005 (non-pooled variances, $df = 6$ and $df = 3$, respectively).

class, while for pre-determined segments, value differences are expected, since the mean map will be averaged from individual maps with different configurations and hence is expected to tend to a flatter 'landscape.' The data in Table II include all adaptively determined segments, from one arbitrarily selected subject, which were 594 msec or longer and belonged to segment classes which occurred at least 3 times; also included in Table II are size-matched sets of pre-determined segments as controls. Table II shows that all 3 expectations are met. In summary, the maps of adaptively determined segments show significantly more distinct spatial characteristics than those of pre-determined segments.

The results in Table II were computed from the complete data sets, i.e., from all sampling time points and from all locations, but used the rarified data set as criterion for adaptive segmentation. Therefore, these results support the notion that the description of the maps at maximal field power times by the peak and trough locations permits

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

recognition of the boundaries of spatially stationary data epochs.

Even though conventional averaging of pre-determined segments results in flattened average maps of power, it obviously does not produce structureless power maps. There is still appreciable variance of band power over space (Table IIC) as can also be seen in the now popular power-mapping systems. This is explained by the relative preponderance of occurrence of few preferred field classes; most frequent are right-posterior-to-anterior types, to a lesser degree left-posterior-to-anterior types (Figs. 4 and 10; see also Lehmann 1971 Fig. 9, 1981 Fig. 18).

Using alpha bandpassed white noise (random over space and time) as input to the segmentation, the mean segment duration (109 msec, S.D. = 76 msec, 1098 segments/2 min) was significantly shorter (t test, $P < 0.005$) than the mean segment duration obtained in our subjects (210 msec). The white noise segments were even significantly shorter ($P < 0.025$) than those of subject 6 who showed the shortest mean duration. Contrary to the real data, all possible 120 classes occurred in the white noise segmentation and, as expected, the white noise data showed a uniform distribution of extrema over space.

Discussion

The adaptively determined segments of the map series belong to different classes as defined by their spatial characteristics ('landscapes'). There appears to be no obvious, simple rule for the succession of the different segment classes in the results in Fig. 8 and in other examined examples of segment sequences, but comprehensive studies still need to be done. However, from a psychophysiological viewpoint we do not expect systematic sequences since we hypothesize that the different micro states or segment classes are installed by the brain as a consequence of the type and content of the information which is momentarily treated (see Koukkou and Lehmann 1983). From this viewpoint, a crucial factor for the prediction of the next state is the individual-specific importance of the momentarily treated information.

Interesting formal aspects of state changes are offered by Wright et al. (1985) who discuss non-linear transitions between linear states, and by Haken's (1983, p. 264) approach of synergetics, which shows that changing a control parameter causes self-organizing systems to pass through several instabilities, i.e., to form a hierarchy of spatial patterns, and that different patterns occur after each identical parameter change, suggesting that different sequences of brain micro states are triggered by identical single input events.

We used time- and space-reduced data for the segmentation of the EEG map series. Segment boundaries of spatially reduced (feature-extracted) and non-reduced data can be compared systematically by using global dissimilarity (formula 2 in Appendix) on segment series of complete maps, employing voltage ranks instead of absolute voltages to concentrate on configuration and to avoid effects of magnitude; global dissimilarity is computed for successive pairs of maps at times of maximal field power; each dissimilarity computation is done twice, using the two maps as they are, and using one map reversed in polarity (equivalent to rectification); the lower of the two dissimilarity values is accepted, since no assumptions are made about the succession of configurations; after 3-point smoothing to repress spurious effects, time points of maximal values in the curve of dissimilarity between successive maps are accepted as segment boundaries. Agreement between this global dissimilarity procedure and our extrema-location method can be tested with the chi-square statistic for comparing two simultaneous neural spike trains (Gerstein et al. 1978). For instance, the data in Fig. 8 which consist of 90 maps at maximal global field power times yielded 17 segment boundaries with the dissimilarity procedure and 19 boundaries with the extrema-location method; 9 coincided at identical time points; goodness of correspondence of the segment boundaries was significant with $\chi^2 = 12.75$, $df = 1$, $P < 0.001$.

Non-linearly interpolated map values (e.g., Ashida et al. 1979; Coppola et al. 1982) would offer the possibility to use locations of extrema

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

which are not restricted to electrode positions; in this case, circular spatial windows of finely graded magnitude could be used, but a very large number of possible segment classes would result, which would require clustering procedures even for a first survey of results. Physiologically meaningful sizes of the spatial windows might be derived from knowledge about the spatial frequencies which exist in momentary maps (but at present no such systematic knowledge is available).

In our spatial approach to adaptive segmentation as well as in other, time-oriented, approaches, the periodic polarity reversals of the electrical fields are accepted as a basic property of a stable brain state. The functional significance of these EEG polarity reversals is still unclear; only a small effect on reaction time is known (Callaway 1962; Dustman and Beck 1965). It might nevertheless be hypothesized that similar brain field configurations of reversed polarity are produced by spatially different neural populations, and hence might show unequal spatial voltage gradients. We tested this in one of our subjects, examining the maps at the times of maximal global field power of the 6 segment classes which each had more than 13 such maps; the segment class with the largest number of cases had 93 maps. There was no significant difference of the gradient magnitude between maps of opposite polarity in any of the 6 classes. This supports the concept that a segment is represented by a spatially stable generator process of periodically reversing polarity.

This paper reports alpha bandpassed data from awake subjects who showed appreciable alpha power during the analysis time. Full band adaptive segmentation produced related but, expectedly, not identical results. Global field power curves for full band are somewhat more noisy, with occasional dominant waves slower than 20/sec (examples in Lehmann 1971). For adaptive segmentation of full band multichannel EEG data from grossly different states such as sleep and wakefulness one might employ a two step approach: first, a time-oriented procedure to segment adaptively the single, global field power curve of the multichannel data into temporally stationary segments in the range of seconds; and secondly, our space-oriented procedure for adaptive segmentation of these larger 'frequency-defined' segments into spatially defined micro states.

The minimal segment duration in our space-based procedure is about a half wave of dominant temporal frequency of the conventional wave shapes (about 50 msec for alpha rhythm) and thus is considerably shorter than in time-oriented adaptive segmentations where minimal times of about a second or more are required (e.g., Creutzfeldt et al. 1985). In addition, and contrary to time-oriented segmentations, the space-oriented procedure can classify a momentary single-map state. (Either the momentary map or the map at the time of the closest maximal value of the global field power curve might be used.) This possibility to classify momentary single maps might permit the recognition of brief state discontinuities such as epileptic events whose momentary spatial maps (Lehmann 1972) differ greatly from normal map configurations.

The functional significance of the segment classes cannot be deduced from their formal aspects. Viewing a segment as a manifestation of an ongoing brain processing step or processing mode, one might speculate that the duration of individual 'thought packages' or 'atoms of mind' are indicated by segment durations. Possibly, the putative time for conscious experience (Libet 1982) may be operationalized for spontaneous thoughts by EEG segmentation. It is conceptually important that about 25% of the analysis time in our data was covered by segments which lasted longer than 500 msec; this is the minimal duration of stimulation for conscious experience if stimuli are applied at an intensity below which no sensation at all is possible (Libet 1982). Related time spans have been reported for switching of attention (Reeves and Sperling 1986: 400 msec), for effective times of leading visual masks (DiLollo 1980, Fig. 7: over 160 msec), and for maximal interstimulus intervals compatible with continuous stereo percepts (W. Skrandies, pers. commun.: 250 msec) and with backward masking of visual targets (Michaels and Turvey 1979, Fig. 13: over 200 msec). This also suggests that functions which are manifest in early components of evoked potentials might not qualify for conscious experience because of their short durations.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The idealistic goal of adaptive time segmentation of brain electrical fields is the recognition of homogeneous basic building blocks of brain information processing, the micro states which constitute the 'atoms' of brain activity. The momentary brain micro state as global entity is conceived to consist of the states of the numerous parallel, automatic brain processes and of the state of the brain's limited capacity channel of controlled processing (consciousness). Mode of information processing (possibly related to 'carrier' and including 'housekeeping' conditions), step of information processing and information content, however, have to be considered in such concepts, since expectedly all of these are manifest in the momentary brain electrical field. This expectation can be extrapolated from studies which examined EEG epochs in the range of a few seconds and reported different EEG characteristics for many psychological measures (e.g., Berkhout et al. 1969; Koukkou and Lehmann 1976; Antrobus 1978; Ehrlichman and Wiener 1980; Lehmann et al. 1981), even though very few measures were examined in any one given study.

We would like to propose the adaptively determined segments of stationary spatial character as a unifying concept for event-related and spontaneous EEG analysis. Space-oriented adaptive segmentation of event-related potential data as a polarity-considering version (Lehmann and Skrandies 1984) of the present strategy has been applied successfully to questions in cognitive psychophysiology (Brandeis and Lehmann 1986). Certain classes of spontaneously occurring segments might have a map configuration which is similar to that of event-related segments ('components') of known significance (example Fig. 16 in Lehmann and Skrandies 1984) and thus might be ascribed a putative, testable functional significance during ongoing EEG activity. In addition, we expect that the study of micro states leads to a deepened understanding of EEG phenomenology.

More information on the functional significance of the different micro states which are manifest as classes of EEG segments can be expected from various types of experiments. Possible physiological variables include event-related potential data whose differing configurations in response to repetitive stimuli even during short recording epochs (Zerlin and Davis 1967) might depend on the momentary map-defined micro state at stimulus arrival, as suggested by single-channel results of time series-defined segments (e.g., Gath et al. 1985; Halász et al. 1985). As to micro state-related behavioral variables, we found in a recent (unpublished) study on 8 normal subjects significant differences in selective motor reaction time depending on the momentary segment class at stimulus delivery: the subjects had to respond to the rare, high pitch tone stimuli which at random replaced some of the lower pitch tone stimuli which were presented in a regular series ('oddball paradigm'), during spontaneously varying segment classes; each subject responsed to 337 rare stimuli; over subjects, there were similar differences in reaction time for the 26 different segment classes which occurred in all subjects (Friedman ANOVA χ^2 40.6, $df = 25$, $P < 0.025$). Another behavioral variable might be different types of private, spontaneous experiences, which trained subjects can report as brief statements upon a signal presented in different micro states. The speculation be permitted that psychotic disturbances of brain functioning might be manifest as an aberrant mosaic of the micro states, i.e., as aberrant occurrence frequency or sequence or duration of the different segment classes.

Appendix

Global field power of a map

Compute the N-weighted standard deviation of the voltages measured at all electrodes vs. one of the N electrodes (Lehmann and Skrandies 1980):

$$\text{global field power} = \left[\frac{1}{n}\sum_{i=1}^{n}(u_i - \bar{u})^2\right]^{0.5}$$

(formula 1)

where n is the number of electrodes on the scalp, u_i is the voltage at electrode i, and $\bar{u}$ is the mean voltage of the n electrodes (this mean voltage is the virtual, so-called average reference).

The value is zero for a completely flat field, and rises with increasing hilliness or relief of the field.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Global dissimilarity of maps

Compute the standard deviation of the voltages at each electrode over all maps (= compute mean map and the S.D.s at all electrodes) using the average reference; sum the standard deviations over all electrodes; divide by the number of electrodes (Lehmann and Skrandies 1980):

$$\text{global dissimilarity} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{1}{m}\sum_{j=1}^{m}(v_{ij}-\bar{v}_i)^2\right]^{0.5}$$

(formula 2)

where n = the number of electrodes, m = the number of maps, v_{ij} is the voltage at electrode i vs. the average reference in map j, and $\bar{v}_i = \frac{1}{m}\sum_{j=1}^{m} v_{ij}$ (the mean voltage over all m maps at electrode i vs. the average reference).

(For m = 2 maps, simply compute the root of the sum (divided by 2) of the squared voltage differences (using the average reference) over all electrodes; divide by the number of electrodes).

If dissimilarity of configuration between maps is to be assessed excluding influences of relief magnitude, the maps must first be scaled to unity global field power, or ranks of voltages must be used.

The value is zero for completely identical maps, and rises with increasing dissimilarity between maps.

The authors wish to thank D. Brandeis, M. Koukkou, G. Ponte, W. Skrandies and two anonymous reviewers for helpful criticism, G. Simon for advice on digital filtering, and R. Müller for collaboration during the data collection.

References

Antrobus, J.S. Dreaming for cognition. In: A.M. Arkin, J.S. Antrobus and S.J. Ellman (Eds.), The Mind in Sleep. Lawrence Erlbaum, Hillsdale, NJ, 1978.

Ashby, W.R. Design for a Brain. Wiley, London, 1960.

Ashida, H., Tatsuno, J., Okamoto, J. and Maru, E. Study of topographic display of EEG with special reference to spectral encephalography. Clin. Electroenceph., 1979, 21: 482–489.

Barlow, J.S. Methods of analysis of nonstationary EEGs, with emphasis on segmentation techniques: a comparative review. J. clin. Neurophysiol., 1985, 2: 267–304.

Barlow, J.S., Creutzfeldt, O.D., Michael, D., Houchin, J. and Epelbaum, H. Automatic adaptive segmentation of clinical EEGs. Electroenceph. clin. Neurophysiol., 1981, 51: 512–525.

Berkhout, J., Walter, D.O. and Adey, W.R. Alterations of human electroencephalogram induced by stressful verbal activity. Electroenceph. clin. Neurophysiol., 1969, 27: 457–469.

Bertrand, O., Perrin, F. and Pernier, J. A theoretical justification of the average reference in topographic evoked potential studies. Electroenceph. clin. Neurophysiol., 1985, 62: 462–464.

Bohdanecky, Z., Bozkov, V. and Radil, T. Acoustic stimulus threshold related to EEG alpha and non-alpha epochs. Int. J. Psychophysiol., 1984, 2: 63–66.

Brandeis, D. and Lehmann, D. Event-related potentials of the brain and cognitive processes: approaches and applications. Neuropsychologia, 1986, 24: 151–168.

Callaway, E. Factors influencing the relationship between alpha activity and visual reaction time. Electroenceph. clin. Neurophysiol., 1962, 14: 674–682.

Coppola, R., Buchsbaum, M. and Rigal, F. Computer generation of surface distribution maps of measure of brain activity. Comput. Biol. Med., 1982, 12: 191–199.

Creutzfeldt, O.D., Bodenstein, G. and Barlow, J.S. Computerized EEG pattern classification by adaptive segmentation and probability density function classification. Clinical evaluation. Electroenceph. clin. Neurophysiol., 1985, 60: 373–393.

DiLollo, V. Temporal integration in visual memory. J. exp. Psychol., Gen., 1980, 109: 75–97.

Dixon, W.J. BMDP Statistical Software. University of California Press, Los Angeles, CA, 1981.

Dustman, R.E. and Beck, E.C. Phase of alpha brain waves, reaction time and visually evoked potentials. Electroenceph. clin. Neurophysiol., 1965, 18: 433–440.

Ehrlichman, H. and Wiener, M.S. EEG asymmetry during covert mental activity. Psychophysiology, 1980, 17: 228–235.

Gath, I., Lehmann, D. and Bar-On, E. Fuzzy clustering of EEG signal and vigilance performance. Int. J. Neurosci., 1983, 20: 303–312.

Gath, I., Bar-On, E. and Lehmann, D. Automatic classification of visual evoked responses. Comp. Meth. Prog. Biomed., 1985, 20: 17–22.

Gerstein, G.L., Perkel, D.H. and Subramanian, K.N. Identification of functionally related neuronal assemblies. Brain Res., 1978, 140: 43–62.

Grandori, F. Dipole localization methods (DLM) and auditory evoked brainstem potentials. Rev. Laryngol. 1984, 105 (Suppl.): 171–178.

Haken, H. Advanced Synergetics, Springer, Heidelberg, 1983.

Halász, P., Pal, I. and Rajna, P. K-complex formation of the EEG in sleep. Acta physiol. hung., 1985, 65: 3–35.

Jansen, B.H., Bourne, J.R. and Ward, J.W. Identification and labeling of EEG graphic elements using autoregressive spectral estimates. Comput. Biol. Med., 1982, 12: 97–106.

Katada, A., Ozaki, H., Suzuki, H. and Suhara, K. Developmental characteristics of normal and mentally retarded children's EEG. Electroenceph. clin. Neurophysiol., 1981, 52: 192–201.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Keesey, U.T. and Nichols, D.J. Fluctuations in target visibility as related to the occurrence of the alpha component of the encephalogram. Vision Res., 1967, 7: 859–879.

Koukkou, M. and Lehmann, D. Human EEG spectra before and during cannabis hallucinations. Biol. Psychiat., 1976, 11: 663–677.

Koukkou, M. and Lehmann, D. Dreaming: the functional state-shift hypothesis. A neuropsychophysiological model. Brit. J. Psychiat., 1983, 142: 221–231.

Koukkou, M., Lehmann, D. and Angst, J. Functional States of the Brain: their Determinants. Elsevier, Amsterdam, 1980.

Lehmann, D. Multichannel topography of human alpha EEG fields. Electroenceph. clin. Neurophysiol., 1971, 31: 439–449.

Lehmann, D. Human scalp EEG fields: evoked, alpha, sleep and spike wave patterns. In: H.H. Petsche and M.A.B. Brazier (Eds.), Synchronization of EEG Activity in Epilepsies. Springer, Vienna, 1972: 307–326.

Lehmann, D. EEG phase differences and their physiological significance in scalp field studies. In: E. Dolce and H. Künkel (Eds.), Computerized EEG Analysis — CEAN. Fischer, Stuttgart, 1975: 102–110.

Lehmann, D. Spatial analysis of evoked and spontaneous EEG potentials. In: N. Yamaguchi and K. Fujisawa (Eds.), Recent Advances in EEG and EMG Data Processing. Elsevier, Amsterdam, 1981: 117–132.

Lehmann, D. EEG assessment of brain activity: spatial aspects, segmentation and imaging. Int. J. Psychophysiol., 1984, 1: 267–276.

Lehmann, D. and Skrandies, W. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroenceph. clin. Neurophysiol., 1980, 48: 609–621.

Lehmann, D. and Skrandies, W. Spatial analysis of evoked potentials in man — a review. Progr. Neurobiol., 1984, 23: 227–250.

Lehmann, D., Beeler, G.W. and Fender, D.H. Changes in patterns of the human electroencephalogram during fluctuations of perception of stabilized retinal images. Electroenceph. clin. Neurophysiol., 1965, 19: 336–343.

Lehmann, D., Koukkou, M. and Andreae, A. Classes of day-dream mentation and EEG power spectra. Sleep Res., 1981, 10: 152.

Lehmann, D., Ozaki, H. and Pal, I. Averaging of spectral power and phase via vector diagram best fits without reference electrode or reference channel. Electroenceph. clin. Neurophysiol., 1986, 64: 350–363.

Libet, B. Brain stimulation in the study of neuronal functions for conscious experience. Hum. Neurobiol., 1982, 1: 235–242.

Lopes da Silva, F.H. Analysis of EEG ongoing activity: rhythms and nonstationarities. In: N. Yamaguchi and K. Fujisawa (Eds.), Recent Advances in EEG and EMG Data Processing. Elsevier, Amsterdam, 1981: 95–115.

McClellan, J.H., Parks, T.W. and Rabiner, L.R. A computer program for designing optimum FIR linear phase filters. IEEE Trans. Audio Electroacoust., 1973, AU-21: 506–526.

Michaels, C.F. and Turvey, M.T. Central sources of visual masking: indexing structures supporting seeing at a single, brief glance. Psychol. Res., 1979, 41: 1–61.

Ohman, A. The orienting response, attention, and learning. In: H.D. Kimmel, E.H. Von Olst and J.F. Orlebeke (Eds.), Drug-Discrimination and State Dependent Learning. Academic Press, New York, 1979: 283–318.

Reeves, A. and Sperling, G. Attention gating in short-term visual memory. Psychol. Rev., 1986, 93: 180–206.

Rohrbaugh, J.W. The orienting reflex: performance and central nervous system manifestations. In: R. Parasuraman and D.R. Davies (Eds.), Varieties of Attention. Academic Press, New York, 1984.

Sidman, R.D., Giambalvo, V., Allison, T. and Bergey, P. A method for localization of sources of human cerebral potentials evoked by sensory stimuli. Sens. Proc., 1978, 2: 116–129.

Walter, D.O., Etevenon, P., Pidoux, B., Tortrat, D. and Guillou, S. Computerized topo-EEG spectral maps: difficulties and perspectives. Neuropsychobiology, 1984, 11: 264–272.

Williams, H.L., Morlock, H.C. and Morlock, J.J. Instrumental behavior during sleep. Psychophysiology, 1966, 2: 208–215.

Woodworth, R.S. and Schlosberg, H. Experimental Psychology. Holt, New York, 1954.

Wright, J.J., Kydd, R.R. and Lees, G.L. State-changes in the brain viewed as linear steady states and non-linear transitions between steady states. Biol. Cybernet., 1985, 53: 11–17.

Zerlin, S. and Davis, H. The variability of single evoked vertex potentials in man. Electroenceph. clin. Neurophysiol., 1967, 23: 468–472.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

VI PHYSICS

The papers in this section provide plausibility arguments for a number of potential theoretical models of anomalous mental phenomena.

The number that appears in the upper right–hand corner of the first page for each publication is keyed to the following descriptions:

25. Tipler, F. J., "Rotating Cylinders and the Possibility of Global Causality Violation," *Physical Review D*, Vol. 9, No. 8, pp. 2203-2206, (April, 1974). The precognition data suggests that causality must be stochastic to avoid the classic problems such as redundant histories. This paper is a technical general relativity argument that shows that under extreme circumstances, causality may be violated.
26. Morris, M. S., Thorne, K. S., and Yurtsever, U., "Wormholes, TIme Machines, and the Weak Energy Condition," *Physical Review Letters,* Vol. 61, No. 13, pp. 1446-1449, (September, 1988). Whereas paper number 25 suggests extreme conditions for possible causality violation, this one demonstrates that time travel may be possible under relatively normal conditions. Morris *et. al.* find they must allow casualty to be stochastic to maintain the validity of the Einstein field equations.
27. Redmount, I., "Wormholes, Time Travel and Quantum Gravity," *New Scientist,* p. 57, (April, 1990). In a non-technical article, Redmount describes how current relativity models might allow for *realizable* time travel. He, too, must invoke a statistical nature for causality.
28. Misra, B. and Sudarshan, E. C. G., "The Zeno's Paradox in Quantum Theory," *Journal of Mathematical Physics,* Vol. 18, No. 4, pp.756-763, (April, 1977). In an important paper in the modern development of quantum theory, Misra and Sudarshan point out that under certain circumstances, physical systems might behave considerable different under "observation" than they do when unobserved. This physics speculation contains implications for anomalous perturbation.
29. Aharonov, Y. and Vardi, M., "Meaning of an Individual 'Feynman Path,'" *Physical Review D,* Vol. 21, No. 8, pp. 2235-2240, (April, 1980). This theoretical paper extends the concepts from Misra and Sudarshan by showing that under *continuous* observations a physical system conforms to the implications of the measuring system rather than to the physical laws usually governing its behavior.
30. Aharonov, Y. and Bohm, D., "Significance of Electromagnetic Potentials in the Quantum Theory," *The Physical Review,* Second Series, Vol. 115, No. 3, pp. 485-491, (August, 1959). One promising possible approach to understanding the transmission characteristics of anomalous mental phenomena is the vector and scalar potentials of the standard electromagnetic theory. This is the initial paper that theoretically showed that these special potentials are not just mathematical improvements, but rather are physically measurable.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

PHYSICAL REVIEW D VOLUME 9, NUMBER 8 15 APRIL 1974

Rotating cylinders and the possibility of global causality violation *

Frank J. Tipler

Department of Physics and Astronomy, University of Maryland, College Park, Maryland 20742

(Received 6 November 1973)

In 1936 van Stockum solved the Einstein equations $G_{\mu\nu} = -8\pi T_{\mu\nu}$ for the gravitational field of a rapidly rotating infinite cylinder. It is shown that such a field violates causality, in the sense that it allows a closed timelike line to connect any two events in spacetime. This suggests that a finite rotating cylinder would also act as a time machine.

Since the work of Hawking and Penrose,[1] it has become accepted that classical general relativity predicts some sort of pathological behavior. However, the exact nature of the pathology is under intense debate at present, primarily because solutions to the field equations can be found which exhibit virtually any type of bizarre behavior.[2,3] It is thus of utmost importance to know what types of pathologies might be expected to occur in actual physical situations. One of these pathologies is causality violation, and in this paper I shall argue that if we make the assumptions concerning the behavior of matter and manifold usual in general relativity, then it should be possible in principle to set up an experiment in which this particular pathology could be observed.

Because general relativity is a local theory with no *a priori* restrictions on the global topology, causality violation can be introduced into solutions quite easily by injudicious choices of topology; for example, we could assume that the timelike coordinate in the Minkowski metric is periodic, or we could make wormhole identifications in Reissner-Nordström space.[4] In both of these cases the causality violation takes the form of closed timelike lines (CTL) which are not homotopic to zero, and these need cause no worries since they can be removed by reinterpreting the metric in a covering space (following Carter,[5] CTL removable by such means will be called trivial—others will be called nontrivial).

In 1949, however, Gödel[6] discovered a solution to the field equations with nonzero cosmological constant that contained nontrivial CTL. Still, it could be argued that the Gödel solution is without physical significance, since it corresponds to a rotating, stationary cosmology, whereas the actual universe is expanding and apparently nonrotating.

The low-angular-momentum Kerr field, on the other hand, cannot be claimed to be without physical relevance: It appears to be the unique final state of gravitational collapse,[7] and so Kerr black holes probably exist somewhere, possibly in the center of our galaxy.[8] This field also contains nontrivial CTL, though the region of causality violation is confined within an event horizon; causality violation from this source could never be observed by terrestrial physicists.[9] In addition, since the CTL must thread their way through a region near the singularity, it is quite possible that matter of a collapsing star will replace this region, as matter replaces the past horizon in the case of spherical collapse.[10] The final Kerr field with collapsed star could be causally well behaved, so the CTL pathology might still be eliminated from general relativity's physical solutions.

I doubt this, because nontrivial causality violation also occurs in the field generated by a rapidly rotating infinite cylinder.

The field of such a cylinder in which the centrifugal forces are balanced by gravitational attraction was discovered by van Stockum in 1936.[11] The metric is expressed in Weyl-Papapetrou form:

$$ds^2 = H(dr^2 + dz^2) + L d\varphi^2 + 2M d\varphi dt - F dt^2, \quad (1)$$

where z measures distance along the cylinder axis, r is the radial distance from the axis, φ is the angle coordinate, and t is required to be timelike at $r = 0$. ($-\infty < z < \infty$, $0 < r < \infty$, $0 \leq \varphi \leq 2\pi$, $-\infty < t < \infty$.) The metric tensor is a function of r alone, and the coordinate condition $FL + M^2 = r^2$ has been imposed (units $G = c = 1$).

It is clear that since $g = \det g_{\mu\nu} = -r^2 H^2$ is negative, the metric signature is $(+ + + -)$ for all $r > 0$, provided $H \neq 0$. van Stockum assumes the Einstein equations

$$G^\mu{}_\nu = -8\pi T^\mu{}_\nu = -8\pi\rho \frac{dx^\mu}{ds} \frac{dx_\nu}{ds},$$

where ρ is the particle mass density. Also

$$\frac{dr}{ds} = \frac{dz}{ds} = 0,$$

$$\frac{d\varphi}{ds} \Big/ \frac{dt}{ds} = \text{constant},$$

$$T = T^\mu{}_\mu = -\rho$$

Copyright © 1974 by The American Physical Society.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

(particle paths required to be timelike).

In a frame in which the matter is at rest, the equations give for the interior field

$$H=e^{-a^2r^2}, \quad L=r^2(1-a^2r^2), \quad \rho=4a^2e^{a^2r^2}, \quad M=ar^2, \quad F=1, \tag{2}$$

where a is the angular velocity of the cylinder. For $r>1/a$, the lines $r=$ constant, $t=$ constant, $z=$ constant are CTL (in fact, by a theorem due to Carter,[5] nontrivial CTL can be found which intersect any two events in the manifold), but one could hope that the causality violation could be eliminated by requiring the boundary of the cylinder to be at $r=R<1/a$. Here the interior solution would be joined to an exterior solution which would be (hopefully) causally well behaved; indeed, the resulting upper bound to the "velocity" aR would equal 1, the speed of light in our units (though the orbits of the particles creating the field are timelike for all r).

van Stockum has developed a procedure which generates an exterior solution for all $aR>0$. When $0<aR<\frac{1}{2}$, the exterior solution is

$$H=e^{-a^2R^2}(r/R)^{-2a^2R^2},$$
$$L=\frac{Rr\sinh(3\epsilon+\theta)}{2\sinh 2\epsilon\cosh\epsilon},$$
$$M=\frac{r\sinh(\epsilon+\theta)}{\sinh 2\epsilon}, \tag{3a}$$
$$F=\frac{r\sinh(\epsilon-\theta)}{R\sinh\epsilon},$$

with

$$\theta=(1-4a^2R^2)^{1/2}\ln(r/R),$$
$$\epsilon=\tanh^{-1}(1-4a^2R^2)^{1/2}.$$

For $aR=\frac{1}{2}$,

$$H=e^{-1/4}(r/R)^{-1/2},$$
$$L=\tfrac{1}{4}Rr[3+\ln(r/R)],$$
$$M=\tfrac{1}{2}r[1+\ln(r/R)], \tag{3b}$$
$$F=(r/R)[1-\ln(r/R)].$$

For $aR>\frac{1}{2}$,

$$H=e^{-a^2R^2}(r/R)^{-2a^2R^2},$$
$$L=\frac{Rr\sin(3\beta+\gamma)}{2\sin 2\beta\cos\beta},$$
$$M=\frac{r\sin(\beta+\gamma)}{\sin 2\beta}, \tag{3c}$$
$$F=\frac{r\sin(\beta-\gamma)}{R\sin\beta},$$

with

$$\gamma=(4a^2R^2-1)^{1/2}\ln(r/R),$$
$$\beta=\tan^{-1}(4a^2R^2-1)^{1/2}$$

[as in the interior solution, $FL+M^2=r^2$, so the metric signature is $(+++-)$ for $R\le r<\infty$].

We see that causality violation is avoided for $aR\le\frac{1}{2}$, but Carter's theorem tells us that it is possible to connect any two events by nontrivial CTL when $aR>\frac{1}{2}$.

There are several objections to be met before this result can be interpreted physically. First c all, Eqs. (3), which van Stockum derived by assuming a special functional form for the $g_{\mu\nu}$, might not be the only candidates for the exterior field; it is known, for instance, that the gravitational field (3a) is *static*[12] in the sense that a "transformation" of the form

$$t'=At+B\varphi, \quad A,B,C,D \text{ constants}$$
$$\varphi'=Ct+D\varphi \tag{4}$$

will eliminate the $g_{\varphi t}$ component. [Transformati is placed in quotes since t' is a periodic coordinate: $t'\equiv t'+B2\pi$. Interpreted globally, the new metric covers a manifold with topology

$$S^2\times(\text{half plane}).$$

We can return to the original topology by taking a covering space, an operation which is *not* equivalent to changing a coordinate system.]

Fortunately, it is easy to prove that (3) are the only possible exterior fields for a rotating infini cylinder. Levy and Robinson[13] have shown that i this case, the Weyl-Papapetrou metric can be written [modulo (4)] in the form

$$ds^2=-e^{2u}(dt+ad\varphi)^2+e^{2(k-u)}(dr^2+dz^2)+r^2e^{-2u}d\varphi^2, \tag{}$$

where u, a, k are functions of r only. A procedur developed by Davies and Caplan[14] and myself allows the equations $R_{\mu\nu}=0$ to be integrated; the s lutions are equivalent to (3). (Details of the uniqueness proof can be found in the Appendix.)

Since the causality problems come from the si nusoid factors of (3c), we might hope to avoid these factors by "transforming" (3a) via (4) and then attempting to join the interior field to the "new" (topologically distinct) field. But the "transformation" (4) will not change the exponer of r, which for $aR>\frac{1}{2}$ become imaginary—in fac for $aR>\frac{1}{2}$, (3a) *is* (3c) with the substitutions $\epsilon=$ and $\theta=i\gamma$.

Thus we expect causality violation to occur in the matter-free space surrounding a rapidly rot ing infinite cylinder. As Thorne[15] has emphasiz however, it is risky to claim that the properties of such a cylinder also hold for realistic cylind

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

In addition to the already mentioned static nature of the field, there is the fact that it is not even asymptotically Minkowskian (especially when $aR > \frac{1}{2}$). Still, the gravitational potential of the cylinder's Newtonian analog also diverges at radial infinity, yet this potential is a good approximation near the surface in the middle of a long but finite cylinder, and if we shrink the rotating cylinder down to a "ring" singularity, we end up with the Kerr field, which also has CTL. These facts suggest that there is a region near the surface of a *finite* cylinder where $g_{\varphi\varphi}$ becomes negative, implying causality violation.

Since $H \neq 0$ for $r \neq 0$, there are no event horizons around the infinite cylinder. By analogy with the static case,[16] I expect this to be true for a finite cylinder; if so, then a timelike line from any event in the universe could enter the region where $g_{\varphi\varphi}$ is negative and return to any other event.[17]

In short, general relativity suggests that if we construct a sufficiently large rotating cylinder, we create a time machine.

I would like to thank Dr. D. Schmidt for helpful discussions, and Professor D. R. Brill for reading the manuscript.

APPENDIX: PROOF THAT VAN STOCKUM'S EXTERIOR SOLUTIONS (3) ARE THE ONLY POSSIBLE EXTERIOR FIELDS FOR AN INFINITE ROTATING CYLINDER

Davies and Caplan have shown[14] that the field equations $R_{\mu\nu}=0$ for the Levy-Robinson metric [Eq. (5)] reduce to

$$\frac{d^2u}{dr^2}+\frac{1}{r}\frac{du}{dr}+\frac{e^{4u}}{2r^2}\left(\frac{da}{dr}\right)^2=0, \quad (A1)$$

$$\frac{d^2a}{dr^2}-\frac{1}{r}\frac{da}{dr}+4\frac{da}{dr}\frac{du}{dr}=0, \quad (A2)$$

$$\frac{2}{r}\frac{dk}{dr}-2\left(\frac{du}{dr}\right)^2+\frac{1}{2r^2}e^{4u}\left(\frac{da}{dr}\right)^2=0. \quad (A3)$$

We have three coupled equations for three functions: second order in u, second order in a, first order in k. Thus we expect five arbitrary constants. A general physical solution to the above system will be defined to be a set of functions a, u, k in which the five constants are allowed to assume all real values from $-\infty$ to ∞. I will show that this general solution is given by Eqs. (3a)–(3c).

Equation (A2) can be written

$$r\frac{d}{dr}\left(\frac{1}{r}e^{4u}\frac{da}{dr}\right)=0.$$

Thus $(1/r)e^{4u}da/dr=2\omega$ (where ω is a constant). Substituting this into (A1), we obtain

$$\frac{d^2u}{dr^2}+\frac{1}{r}\frac{du}{dr}+2\omega^2e^{-4u}=0. \quad (A4)$$

Suppose first that $\omega=0$. Then a little manipulation yields

$$u=A(\ln r)+B, \quad k=A^2(\ln r)+C, \quad a=D,$$

where A, B, C, D are constants.

By the transformation $t=t'-a\varphi$, $\varphi=\varphi'$, $z=z'$, $r=r'$, we discover that except for global topology this solution is just the Weyl solution (3a).

Suppose now that $\omega\neq 0$. It is at this point that Davies and Caplan err; their "general" solution in fact places implicit restrictions on the value of their constant A. The complete general solution is obtained via the following procedure. Let $v=e^{-4u}$, $p=(\omega r)^2$, so that $u=-\frac{1}{4}\ln(v)$, and $d/dr=2\omega^2r(d/dp)$, which gives

$$\frac{du}{dr}=-\frac{1}{4v}\frac{dv}{dp}2\omega^2r.$$

Equation (A4) becomes

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{du}{dr}\right)+2\omega^2e^{-4u}$$
$$=\frac{2\omega^2r}{r}\frac{d}{dp}\left(-\frac{2\omega^2r^2}{4v}\frac{dv}{dp}\right)+2\omega^2v=0$$

or

$$\frac{d}{dp}\left(\frac{p}{v}\frac{dv}{dp}\right)-2v=0. \quad (A5)$$

Let $w=pv$, giving $dv/dp\equiv v'=w'/p-w/p^2$. (A5) becomes

$$\frac{d}{dp}\left(\frac{pw'}{w}\right)-\frac{2w}{p}=0. \quad (A6)$$

Let $t=\ln(p)$, $d/dp=(1/p)d/dt$. (A6) becomes

$$\frac{d}{dt}\left(\frac{\dot{w}}{w}\right)-2w=0. \quad (A7)$$

Let $Q=\dot{w}\equiv dw/dt$, $d/dt=(Q)d/dw$. (A7) becomes

$$Q\frac{d}{dw}\left(\frac{Q}{w}\right)-2w=0$$

or

$$\frac{d}{dw}\left(\frac{Q}{w}\right)-2\frac{w}{Q}=0.$$

Thus

$$\frac{Q}{w}d\left(\frac{Q}{w}\right)=2dw$$

or

$$(Q/w)^2=4w\pm A^2,$$

which can be written

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

$$\dot{w} = w(4w \pm A^2)^{1/2}. \qquad (A8)$$

The next integral depends on the sign choice in (A8). First choose the + sign. Then performing the integration, we obtain

$$\frac{1}{A}\ln\left[\frac{(4w+A^2)^{1/2}-A}{(4w+A^2)^{1/2}+A}\right] = \ln p + \frac{1}{A}\ln B.$$

[The constant of integration $(1/A)\ln B$ has values from $-\infty$ to ∞, though $0<B<\infty$.] This can be inverted (after the appropriate substitutions are made) to give

$$u = \frac{1}{4}\ln\left[\frac{(1-\omega^{2A}r^{2A}B)^2}{A^2(\omega r)^{2A-2}B}\right],$$

which is identical to Eq. (2.3) of Davies and Caplan (in Ref. 14). The computation proceeds as they outline to obtain k and a. Frehland[14] has shown that this solution is the same as the Weyl solution (3a).

Suppose now that $A=0$. We get

$$g_{zz} = g_{rr} = Fr^{-1/2},$$
$$g_{\varphi t} = -r(1 - E\omega \ln\omega^2 r^2 D^2),$$
$$g_{tt} = \omega r \ln(\omega^2 r^2 D^2),$$
$$g_{\varphi\varphi} = -rE(2 + E\omega \ln\omega^2 r^2 D^2),$$

where E, D, F are constants. These solutions ar identical to (3b), with a suitable choice of constants. Suppose now that the integration constar is $-A^2$. We obtain

$$g_{tt} = \frac{2\omega r}{A}\cos[\ln(\omega^A r^A)+C],$$
$$g_{zz} = g_{rr} = Fr^{-(1+A^2)/2}, \qquad (A1$$
$$g_{\varphi t} = r\{\sin[\ln(\omega^A r^A)+C] + D\cos[\ln(\omega^A r^A)+C]\}.$$

$g_{\varphi\varphi}$ is determined by the relation $FL + M^2 = r^2$, where A, C, D, F are constants.

Thus the general exterior field is given by (3).

---

*Work supported in part by the National Science Foundation under Grant No. GP-25548.

[1]S. W. Hawking and R. Penrose, Proc. R. Soc. Lond. A 314, 529 (1970).

[2]C. W. Misner, in *Relativity Theory and Astrophysics: Relativity and Cosmology*, edited by J. Ehlers (American Mathematical Society, Providence, R.I., 1967), Vol. 8, p. 160.

[3]R. Geroch, Ann. Phys. (N.Y.) 48, 526 (1968).

[4]B. Carter, Phys. Lett. 21, 423 (1966).

[5]B. Carter, Phys. Rev. 174, 1559 (1968). Carter's causality theorem can be stated as follows: A necessary and sufficient condition for nontrivial causality violation in a connected, time-oriented spacetime with a timewise orthogonally transitive Abelian isometry group is the nonexistence of a covariant vector in the Lie algebra such that the corresponding differential form in the surface of transitivity is everywhere well behaved and everywhere timelike. If the above criterion is satisfied, then there exist both future- and past-directed timelike lines between any two points of the spacetime. For the van Stockum metrics (2) and (3a)–(3b), the group generated by the Killing vectors $(\partial/\partial z, \partial/\partial t, \partial/\partial\varphi)$ is timewise orthogonally transitive and Abelian. It is easily checked that for $r>1/a$ in (2) and $aR>\frac{1}{2}$ in (3), there is no linear combination $\psi = At + B\varphi + Cz$ (where A, B, C are constants) such that the form $d\psi$ is everywhere timelike.

[6]K. Gödel, Rev. Mod. Phys. 21, 447 (1949).

[7]R. M. Wald, Phys. Rev. Lett. 26, 1653 (1971).

[8]C. W. Misner, Phys. Rev. Lett. 28, 994 (1972).

[9]For $a^2+e^2>m^2$ there are no event horizons and so causality violation is global, but it is not clear that a star with such high values of angular momentum and/or charge would collapse sufficiently far to uncover the region where $g_{\varphi\varphi}$ changes sign (see Ref. 7). Penrose has argued [in Proceedings of the Sixth Texas Symposium on Relativistic Astrophysics, 1972 (unpublished)] that a naked Kerr singularity would be a good model for a rapidly rotating star which has collapsed into a disk. CTL would be expected when $e=0$, but one might contend that these occur so close to the singularity (and hence in regions where we expect general relativity to break down anyway) that they are without physical significance. van Stockum's work shows, however, that CTL are not necessarily associated with extreme curvature in physically significant situations.

[10]C. W. Misner, in *Astrophysics and General Relativity* edited by M. Chrétien, S. Deser, and J. Goldstein (Gordon and Breach, New York, 1969), Vol. 1.

[11]W. J. van Stockum, Proc. R. Soc. Edinb. 57, 135 (193 Other authors, such as S. C. Maitra [J. Math. Phys. 7, 1025 (1966)] have noted that the van Stockum interi solution possesses CTL.

[12]J. Ehlers and W. Kundt, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962), p. 84.

[13]H. Levy and W. J. Robinson, Proc. Camb. Philos. Soc. 60, 279 (1964).

[14]H. Davies and T. A. Caplan, Proc. Camb. Philos. Soc. 69, 325 (1971); E. Frehland, Commun. Math. Phys. 23, 127 (1971).

[15]K. S. Thorne, Comments Astrophys. Space Phys. 2, 191 (1970).

[16]W. Israel, Nature 216, 148 (1967); 216, 312 (1967).

[17]Assuming, of course, that the cylinder has existed for all time. If it is created, then this statement will have to be qualified somewhat, but observable causali violation will still occur.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

26

Wormholes, Time Machines, and the Weak Energy Condition

Michael S. Morris, Kip S. Thorne, and Ulvi Yurtsever
Theoretical Astrophysics, California Institute of Technology, Pasadena, California 91125
(Received 21 June 1988)

It is argued that, if the laws of physics permit an advanced civilization to create and maintain a wormhole in space for interstellar travel, then that wormhole can be converted into a time machine with which causality might be violatable. Whether wormholes can be created and maintained entails deep, ill-understood issues about cosmic censorship, quantum gravity, and quantum field theory, including the question of whether field theory enforces an averaged version of the weak energy condition.

PACS numbers: 04.60.+n, 03.70.+k, 04.20.Cv

Normally theoretical physicists ask, "What are the laws of physics?" and/or, "What do those laws predict about the Universe?" In this Letter we ask, instead, "What constraints do the laws of physics place on the activities of an arbitrarily advanced civilization?" This will lead to some intriguing queries about the laws themselves.

We begin by asking whether the laws of physics permit an arbitrarily advanced civilization to construct and maintain wormholes for interstellar travel. Such a wormhole is a short "handle" in the topology of space, which links widely separated regions of the Universe (Fig. 1). The Schwarzschild metric, with an appropriate choice of topology, describes such a wormhole.[1,2] However, the Schwarzschild wormhole's horizon prevents two-way travel, and its throat pinches off so quickly that it cannot be traversed in even one direction.[2,3] To prevent pinchoff (singularities) and horizons, one must thread the throat with nonzero stress and energy.[4] One then faces two questions: (i) Does quantum field theory permit the kind of stress-energy tensor that is required to maintain a two-way-traversible wormhole? (ii) Do the laws of physics permit the creation of wormholes in a universe whose spatial sections initially are simply connected? These questions take on added importance when one recognizes (see below) that, if the laws of physics permit traversible wormholes, then they probably also permit such a wormhole to be transformed into a "time machine" with which causality might be violatable. In the remainder of this Letter we discuss in turn the creation of wormholes, their maintenance by quantum-field-theoretic stress-energy tensors, and their conversion into time machines.

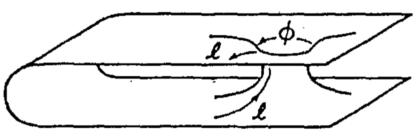


FIG. 1. A wormhole and the external universe at a specific moment of time, embedded in a fictitious higher-dimensional space.

Wormhole creation.—Wormhole creation, with such mild spacetime curvature that classical general relativity is everywhere valid, must be accompanied by closed timelike curves and/or a noncontinuous choice of the future light cone,[5] and also by a violation of the "weak energy condition."[6] Specific spacetimes with such wormhole creation are known.[7] However, it is not known whether the stress-energy tensors required by the Einstein equations in those spacetimes are permitted by quantum field theory.

Wormhole creation accompanied by extremely large spacetime curvatures would be governed by the laws of quantum gravity. A seemingly plausible scenario entails quantum foam[1,8] [finite probability amplitudes for a variety of topologies on length scales of order of the Planck-Wheeler length, $(G\hbar/c^3)^{1/2} = 1.3\times10^{-33}$ cm]: One can imagine an advanced civilization pulling a wormhole out of the quantum foam and enlarging it to classical size. This might be analyzable by techniques now being developed for computation of spontaneous wormhole production by quantum tunneling.[9]

Wormhole maintenance.—For any traversible wormhole a two-sphere surrounding one mouth (but well outside it where spacetime is nearly flat), as seen through the wormhole from the other mouth, is an outer trapped surface. This implies[10] (since there is no event horizon) that the wormhole's stress-energy tensor $T_{\alpha\beta}$ must violate the averaged weak energy condition[11] (AWEC); i.e., passing through the wormhole there must be null geodesics, with tangent vectors $k^\alpha = dx^\alpha/d\zeta$, along which $\int_0^\infty T_{\alpha\beta}k^\alpha k^\beta d\zeta < 0$. (Our conventions are those of Ref. 2.) Thus, if one could show that quantum field theory forbids violations of AWEC, one could rule out advanced civilizations maintaining traversible wormholes.

Specialize to a static, spherical wormhole, with spacetime metric[4]

$$ds^2 = -e^{2\Phi}dt^2 + dl^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2),$$

where Φ and r are functions of proper radial distance l. Set $l=0$ at the throat ($l<0$ on the "left" side of the throat and $l>0$ on the "right" side). Far from the

 © 1988 The American Physical Society
Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

throat, $r \cong |l| - M \ln(|l|/r_0)$ and $\Phi \cong -2M/|l|$, where M is the wormhole's mass. As l increases from $-\infty$ to 0, r decreases monotonically to a minimum value r_0, the throat's "radius"; and as l increases onward to $+\infty$, r increases monotonically. Φ is everywhere finite (no horizons). For such a wormhole, AWEC is violated on radial null geodesics, and its violation can be expressed as $\int_{l_1}^{\infty}(\tau-\rho)e^{-\Phi}dl > 0$ for any $l_1 < 0$. Here ρ and τ are the energy density and radial tension, $\rho = e^{2\Phi} \times T^{tt}$, $\tau = -T^{ll}$, and the affine parameter is $\zeta = \int e^{\Phi} dl$.

The following model explores the use of the "Casimir vacuum"[12] (a quantum state of the electromagnetic field that violates the unaveraged weak energy condition[11]) to support a wormhole: At $l = \pm s/2$, we place two identical, perfectly conducting spherical plates with equal electric charges Q; we require $s \ll r_0$; and we carry out our analysis only up to fractional errors of order s/r_0. Between the plates the electromagnetic field is in the Casimir vacuum state for which ρ, τ, and the lateral pressure $p \equiv T^{\theta}{}_{\theta} = T^{\phi}{}_{\phi}$ are[13]

$$\tau = 3p = -3\rho = (3\pi^2/720)(\hbar/s^4) \equiv \tau_C.$$

Outside the plates is a classical, radial Coulomb field with $\tau = \rho = p = Q^2/8\pi r^4$, which produces a Reissner-Nordstrom spacetime geometry.[2] Force balance at the plates requires $Q = (8\pi r_0^4 \tau_C)^{1/2}$. The Einstein field equations can then be satisfied throughout if[4] (i) $\tau_C = (8\pi r_0^2)^{-1}$, so that $s = (\pi^3/30)^{1/4}(r_0\sqrt{\hbar})^{1/2}$ [$s \sim 10^{-10}$ cm for r_0 = (distance from Sun to Earth)]; (ii) the wormhole's charge Q and mass M, as measured by distant observers, are both equal to r_0 (aside from fractional corrections $\lesssim s/r_0$); and (iii) the energy per unit area σ of each plate is such as to violate AWEC: $2\sigma < \frac{4}{3}\tau_C s$, or more precisely $\frac{4}{3}\tau_C s - 2\sigma > O(\tau_C s^2/r_0)$.

This violation of AWEC is compatible with a total nonnegative energy of plates plus Casimir field, $2\sigma + \rho s = 2\sigma - \frac{1}{3}\tau_C s \geq 0$. However, if quantum field theory requires that the plates' mass-to-charge ratios exceed that of an electron, $4\pi r_0^2 \sigma/Q > m/e$, then $2\sigma < \frac{4}{3}\tau_C s$ corresponds to a plate separation smaller than the electron Compton wavelength,

$$s < [(\pi^3/270)e^2/\hbar]^{1/2}\hbar/m = 0.029\hbar/m,$$

which might well be forbidden. To determine whether $\sigma < \frac{4}{3}\tau_C s$ is allowed would require explicit study of quantum-field-theory models for the plates (a task we have not attempted) or a general theorem that quantum field theory forbids violations of AWEC.

Conversion of wormhole into time machine.— Figure 2 is a spacetime diagram for the conversion of a spherical, traversible wormhole into a time machine. Shown unstippled is the nearly flat spacetime outside the wormhole, with Lorentz coordinates T,Z (shown) and X,Y (suppressed). Shown stippled is the wormhole interior, i.e., the region of large spacetime curvature. The central lines of the stippled strips are the wormhole throat,

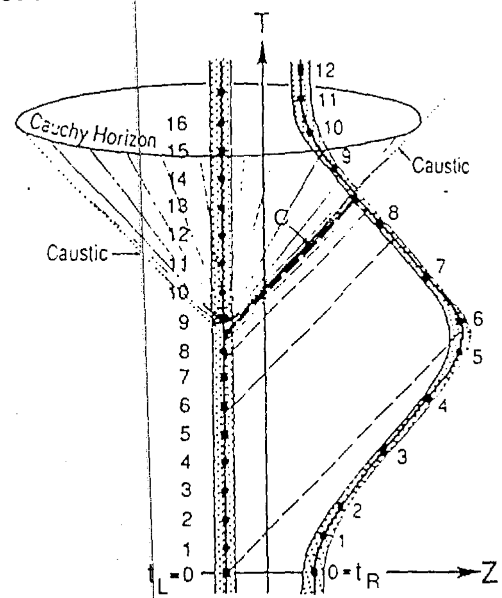


FIG. 2. Spacetime diagram for conversion of a wormhole into a time machine.

parametrized by a time coordinate t introduced below.

At $T = 0$, the wormhole's mouths are at rest near each other. Subsequently, the left mouth remains at rest while the right mouth accelerates to near-light speed, then reverses its motion and returns to its original location. The advanced beings can produce this motion by pulling on the right mouth gravitationally or electrically. This motion causes the right mouth to "age" less than the left as seen from the exterior. Consequently, at late times by traversing the wormhole from right mouth to left, one can travel backward in time (i.e., one can traverse a closed timelike curve) and thereby, perhaps, violate causality.

The metric inside the accelerating wormhole and outside but near its mouths is

$$ds^2 = -(1 + glF\cos\theta)^2 e^{2\Phi}dt^2 + dl^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2).$$

Here $\Phi = \Phi(l)$ and $r = r(l)$ are the same functions as for the original, static wormhole; $F = F(l)$ is a form factor that vanishes in the left half of the wormhole $l \leq 0$, and rises smoothly from 0 to 1 as one moves rightward from the throat to the right mouth; and $g = g(t)$ is the acceleration of the right mouth as measured in its own asymptotic rest frame. Just outside the right and left mouths the transformation from wormhole coordinates to external, Lorentz coordinates (with $ds^2 \cong -dT^2 + dX^2 + dY^2 + dZ^2$) is $T = T_R + v\gamma l\cos\theta$, $Z = Z_R + \gamma l \times \cos\theta$, $X = l\sin\theta\cos\phi$, $Y = l\sin\theta\sin\phi$; $T = t$, $Z = Z_L + l \times \cos\theta$, $X = l\sin\theta\cos\phi$, $Y = l\sin\theta\sin\phi$. Here Z_L is the time-independent Z location of the left mouth's center; $Z = Z_R(t), T = T_R(t)$ is the world line of the right mouth's center with $dt^2 = dT_R^2 - dZ_R^2$; $v(t) \equiv dZ_R/dT_R$ is

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

the velocity of the right mouth; and $\gamma(t) \equiv (1-v^2)^{-1/2}$. The acceleration appearing in the wormhole metric is $g(t) = \gamma^2 dv/dt$. The right mouth's maximum acceleration g_{max} and the distance S through the wormhole from left mouth to right must satisfy $g_{max}S \ll 1$, $g_{max}|dg/dt|^{-1} \gg S$. This guarantees that with tiny fractional changes of $T^{\alpha\beta}$, the wormhole's size and shape are held nearly constant throughout the acceleration.

The region of spacetime containing closed timelike curves is separated from that without such curves by a Cauchy horizon. One might have expected this Cauchy horizon to be unstable (in accord with strong cosmic censorship[14]). Indeed, in the analogous two-dimensional (2D) Misner space[15] [obtainable from 2D Minkowskii spacetime by identification of $(T,Z) = (\xi,0)$ with $(T,Z) = (\gamma\xi, L-\gamma v\xi)$, where $L>0$, $v>0$, $\gamma = (1-v^2)^{-1/2}$, and ξ runs from $-\infty$ to $L/\gamma v$] the Cauchy horizon H [located at $T-Z = L/(B-1)$ where $B \equiv (1+v)^{1/2}/(1-v)^{1/2}$] *is* unstable. Rightward-propagating waves in Misner space get boosted in frequency by a factor B with each passage through the identification world line, and they pass through it infinitely many times as they approach the Cauchy horizon, H. As a result, the stress-energy tensor of such waves diverges at H—presumably thereby preventing the spacetime from evolving the closed timelike curves that it otherwise would have beyond H. [This is the same instability as occurs at the Misner hypersurface in Taub-NUT (Newman-Unti-Tamburino) space.[10,16]]

In our 4D wormhole spacetime the Cauchy horizon H seems not to suffer this particular instability. There H possesses precisely one closed null generator[17]: the curve C in Fig. 2, which runs along the Z axis from left mouth to right mouth, through the throat, then along the Z axis again. The remainder of H consists of null geodesics (very thin lines in Fig. 2) which peel off C to form a future light cone of the wormhole's left mouth (with caustics, where future-directed generators leave the horizon, along the Z axis to the left mouth and right of the right mouth). The most likely place for the Misner-type instability is on C. Indeed, a light ray (dashed curve of Fig. 2), running along the Z axis before horizon formation, gets Doppler shifted by the factor B with each traversal through the wormhole; and it traverses the wormhole infinitely many times as it asymptotes to C. However, the wormhole's AWEC violation causes the throat to act like a diverging lens with focal length $f \simeq r_0/2$. Correspondingly, if D is the Z distance between wormhole mouths as measured along C, waves propagating along the dashed curve get reduced in amplitude by $f/D \simeq r_0/2D$ with each round trip from left mouth to right and through the wormhole. If the advanced beings arrange that $(f/D)B < 1$, the reduction in amplitude will dominate over the boost in frequency, thereby reducing the wave energy with each round trip and leaving the Cauchy horizon immune to this instability. We suspect, in fact, that the Cauchy horizon is fully stable and thus constitutes a counterexample to the conjecture of strong cosmic censorship.[14]

For Misner space[15] (as also for Taub-NUT space[10,16]) the extension of spacetime through the Cauchy horizon H is not unique: In one extension all "leftward" causal geodesics (those with initial rightward velocities less than a critical value) are well behaved, while all "rightward" causal geodesics terminate at H after finite affine parameter; in another extension the rightward geodesics are well behaved and the leftward terminate. By analogy, one might expect there to exist other extensions (besides Fig. 2) of the 4D wormhole spacetime beyond H; one might even hope that in the real universe such a wormhole would actually find and evolve into an extension (possibly nonanalytic) with no closed timelike curves. However, because the set of spacetime geodesics that terminate on H is of "measure zero" (it is a four-parameter set compared to six parameters for generic geodisics), we suspect (provided H is indeed stable) that the extension beyond H is uniquely that of Fig. 2. More generally, we speculate that whenever a spacetime has a fully stable Cauchy horizon, its evolution through that horizon is unique. Similarly, we speculate (as has been suggested to us by Friedman[7]) that in such a spacetime physical fields, both quantum and classical, evolve through and beyond H in unique ways.[18] For example, if initial data for a classical scalar field ψ are specified at $T=0$ in the spacetime of Fig. 2, the resulting evolution via $\Box\psi = 0$ beyond H will exist and be unique. This is because the set of causal geodesics to the future of H, which do not extend back through $T=0$, is of measure zero (is only a four-parameter set); and such geodesics experience an infinity of "diverging-lens" wormhole traversals as one follows them backward in local time to points from which they could carry unspecified initial data. This makes us doubt that any "new" field ψ can propagate into the spacetime along them.

This wormhole spacetime may serve as a useful test bed for ideas about causality, "free will," and the quantum theory of measurement. As an infamous example, can an advanced being measure Schrödinger's cat to be alive at an event P (thereby "collapsing its wave function" onto a "live" state), then go backward in time via the wormhole and kill the cat (collapse its wave function onto a "dead" state) before it reaches P?

For helpful discussions or correspondence, we thank Bruce Allen, Phillip Candelas, S. Chandrasekhar, Bryce DeWitt, Sam Finn, John Friedman, Robert Geroch, Leonard Parker, Don Page, Nicolas Papastamatiou, Thomas Roman, Carl Sagan, and Robert Wald. One of us (K.S.T.) thanks the Departments of Physics and Astronomy and the Phenomenology Group of the University of Wisconsin at Madison for hospitality during this research. This work was supported in part by National Science Foundation Grant No. AST85-14911.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

[1]J. A. Wheeler, *Geometrodynamics* (Academic, New York, 1962).

[2]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).

[3]M. D. Kruskal, Phys. Rev. 119, 1743 (1960).

[4]M. S. Morris and K. S. Thorne, Am. J. Phys. (to be published).

[5]Theorem 2 of R. P. Geroch, J. Math. Phys. 8, 782 (1967); for further discussion, see R. P. Geroch, Ph.D. thesis, Princeton University, 1967 (unpublished).

[6]F. J. Tipler, Phys. Rev. Lett. 37, 879 (1976); Sec. 5 of F. J. Tipler, Ann. Phys. (N.Y.) 108, 1 (1977).

[7]The method for mathematical construction of such spacetimes is given by P. Yodzis, Commun. Math. Phys. 26, 39 (1972). John Friedman and Zhang Huai (private communication) have given a lovely explicit example in which two wormholes are created in a compact region of spacetime.

[8]S. W. Hawking, Phys. Rev. D 37, 904 (1988).

[9]H. Kandrup and P. Mazur, "Particle Creation and Topology Change in Quantum Cosmology" (unpublished).

[10]Proposition 9.2.8 of S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge Univ. Press, Cambridge, 1973), as adapted to the wormhole topology and as modified by replacement of the weak energy condition by the AWEC with an argument from D. Deutsch and P. Candelas, Phys. Rev. D 20, 3063 (1980), and references therein. We thank Don Page for pointing out a variant of this argument.

[11]F. J. Tipler, Phys. Rev. D 17, 2521 (1978); T. A. Roman, Phys. Rev. D 33, 3526 (1986), and 37, 546 (1988).

[12]Deutsch and Candelas, in Ref. 10.

[13]These renormalized stresses and energy density are the same as for flat plates in flat spacetime. The wormhole's spacetime geometry will modify them by fractional amounts $\lesssim s/r_0$, which we ignore. The nonzero extrinsic curvature of the plates, which is of order s/r_0^2, will give rise to renormalized stresses and energy density of order $(hs/r_0^2)(l \pm s/2)^{-3}$ which diverge at the plates; but these can be made negligible by giving the plates a finite skin depth $\delta \sim (s^4/r_0)^{1/3} \ll r_0$; see Ref. 12.

[14]F. J. Tipler, C. J. S. Clarke, and G. F. R. Ellis, in *Einstein Centenary Volume*, edited by A. Held and P. Bergman (Plenum, New York, 1979).

[15]C. W. Misner, in *Relativity Theory and Astrophysics I: Relativity and Cosmology*, edited by J. Ehlers (American Mathematical Society, Providence, RI, 1967), p. 160; also Sec. 5.8 of Ref. 10.

[16]C. W. Misner and A. H. Taub, Zh. Eksp. Teor. Fiz. 55, 233 (1968) [Sov. Phys. JETP 28, 122 (1969)].

[17]For a discussion of Cauchy horizons with structures similar to this, see F. J. Tipler, Ph.D. thesis, University of Maryland, 1976 (unpublished), Chap. 4.

[18]For a discussion of uniqueness of evolution of nonglobally hyperbolic spacetimes from a somewhat different viewpoint, see R. M. Wald, J. Math. Phys. 21, 2802 (1980).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Wormholes, time travel and quantum gravity

"Wormholes" in space and time have suddenly become a focus of attention. Large ones might be used as time machines; small ones may determine the values of all the constants of nature

Ian Redmount

Andrew Kulman

THE PAST few years have seen an explosion of interest in what theoretical physicists call "wormholes". These are tunnels in the geometry of space and time, connecting otherwise distant or completely disconnected regions of the Universe. In fact, there have been two explosions in two almost unrelated subjects. One is macroscopic wormholes, the kind that science fiction writers or sufficiently advanced civilisations might use for space travel across cosmic distances. The other is microscopic wormholes, on a scale 20 orders of magnitude smaller than an atomic nucleus. At this scale, space and time should obey the rules of quantum physics rather than classical laws.

Neither type of wormhole is a new idea. Theorists have known about large-scale wormholes for more than 70 years—since shortly after Albert Einstein put forward the general theory of relativity, which relates gravity to the geometry of space and time. For 30 years, physicists have conjectured that microscopic wormholes might play a crucial role in understanding the structure of elementary particles or in developing a quantum theory of gravity. Recently, however, researchers have found that both types of objects may have some remarkable, previously unsuspected properties. Large-scale wormholes could provide a relatively easy means of travelling backwards in time, with all the potential complications that entails. Microscopic wormholes might, through their contribution to the quantum mechanics of the Universe, determine the values of all the constants in all the laws of physics. Much of this wormhole work is speculative and some very controversial, but that is why these subjects have generated such excitement lately.

Michael Morris and Ulvi Yurtsever, and their PhD thesis adviser Kip Thorne of the California Institute of Technology, began discovering new features of large-scale wormholes in 1985. They were trying to flesh out the idea of using wormholes for interstellar space travel, as described in Carl Sagan's novel *Contact*. What, they asked, do the known laws of physics require for such a thing to work?

A classical, large-scale wormhole is a solution of the field equations of Einstein's general theory of relativity, a geometry of space and time, or "space-time", in which two regions of the Universe are connected by a short, narrow "throat". The best-known such geometry is the spherically symmetrical, matter-free solution discovered by Karl Schwarzschild in 1916. A portion of this solution, omitting one of the exterior regions and the throat, serves to describe the space-time around a spherical, non-rotating star, planet, or other object. A slightly larger portion describes a non-rotating, electrically neutral black hole. But neither of these objects connects distant regions of the Universe. Only the full wormhole geometry does that (see Figure 1).

Figure 1 is deceiving, however. The wormhole shown is not a static structure; it represents the shape of space at the single instant of maximum expansion of the throat. The Schwarzschild wormhole actually expands from zero throat radius to the maximum shown in Figure 1, then shrinks back to zero. It does this so quickly that no traveller, even one moving at the speed of light, can ever pass from one mouth of the hole to the other. Such a wormhole is not "traversable". Any matter falling into the wormhole from the surrounding space hastens the contraction, so that the traveller cannot even come close to making a safe passage.

Other wormhole solutions to Einstein's equations—for an electrically charged or a rotating black hole, for example—while they are ostensibly traversable, suffer from the same problem. Any matter that falls in, or any radiation, is so concentrated and amplified by the gravitational fields of the hole that its own gravity alters the spacetime and closes off the hole. Moreover, all these wormholes exert tidal gravitational forces as strong as those of a black hole of the same size; a hole that is metres or kilometres in size would shred travellers of human dimensions long before they even got near it. Clearly a hole suitable for space travel requires some novel modifications.

What the team from Caltech did was to construct mathematically wormhole geometries that would allow passage, with throats that stayed open and gravitational fields such that travellers encountered only modest accelerations and tidal forces. The equations of general relativity then told them what kinds of matter and energy were needed to make up the holes. They found that the throats of their holes had to be threaded by matter or fields with enormous negative pressure, in other words, the matter would have to have a tension, rather like a stretched spring.

For a hole a kilometre or so across, the size of this tension is similar to the pressure at the centre of a neutron star (a star with about the same mass as our Sun, occupying the volume of a large mountain on Earth). For a smaller hole, the tension would be greater. Most crucially, the magnitude of the tension must be greater than the energy per unit volume (the energy density) of the matter itself.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Matter with this property is called "exotic". In familiar matter, tensions and pressures are always far smaller than the energy density: the breaking tension of steel, for example, is some 12 orders of magnitude (10^{12} times) smaller than its energy density. A tension larger than the energy density implies that some observers—moving rapidly with respect to the wormhole—will observe the matter to have negative energy density. Einstein's general theory of relativity relates the density and pressure, or tension, measured by one set of observers to those of another. The relationship guarantees a positive energy density for all observers if pressure or tension is smaller than energy density for any one observer, but not otherwise. Einstein's field equations imply that any traversable wormhole must contain some form of this exotic matter.

We do not know whether this requirement rules out the possibility of traversable wormholes or not. Physicists usually assume that matter obeys certain energy conditions, among which is the requirement that all observers measure positive energy densities. Situations exist, however, in which these conditions are known to be false.

The electromagnetic field between two metal plates can, for example,, give rise to a negative energy density. Because, according to quantum mechanics, the electric and magnetic fields obey Heisenberg's uncertainty principle, they fluctuate minutely, rather than holding precise, classical values. Even the vacuum contains these field fluctuations. The energy of the fluctuations in the field between conducting plates is actually less than that in the free vacuum; that is, it is negative. This effect is named after the Dutch physicist Hendrik Casimir, who calculated it in 1948, and it has since been confirmed in the laboratory.

The evaporation of black holes, discovered by Stephen Hawking at the University of Cambridge in 1974, also involves negative energy densities.

No one knows whether exotic matter of the density and extent required to make a traversable wormhole can exist or not. If it can, and if it interacts weakly enough with other matter to avoid harming the traveller, or can be confined within the wormhole away from the traveller's path, then traversable wormholes remain a physical possibility.

The results of other theorists support this. Matt Visser of Washington University, St Louis, has found a kind of wormhole that travellers can pass through it without encountering any matter, exotic or otherwise, and without feeling any forces at all. He takes two copies of what is called Minkowski space-time—this is infinite, empty space-time with no matter or gravitational fields—and excises an identical region from each, and joins them at the boundaries of the excised regions. The energy densities and pressures on the boundary surface, now the throat of the wormhole, are specified by Einstein's field equations. If, for example, the junction surface is a cube, then all the exotic matter is confined to "struts" on the edges of the cube. A traveller can pass from one Minkowski region to the other through a face of the cube, untouched by any matter or force. Visser's work also suggests that wormholes like this could be made stable—they would neither collapse nor explode, clearly another requirement for holes usable for travel.

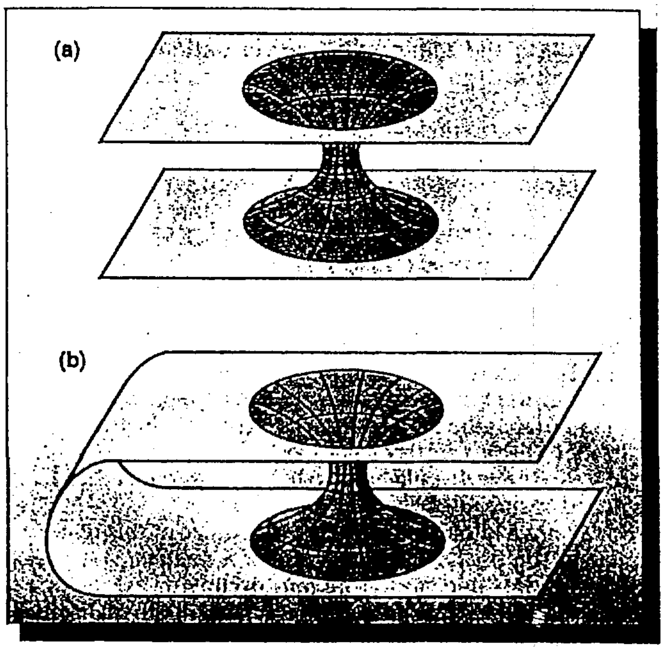


Figure 1 *A wormhole can connect two regions, as in* **a**, *or two points of a single region as in* **b**

In a similar vein, recent work by Ian Moss, Felicity Mellor, and Paul Davies at the University of Newcastle upon Tyne indicates that in our expanding Universe, some wormholes may not be forced to collapse by the effects of infalling matter and radiation. So this may not be the problem for wormholes in our expanding Universe that it is for holes in hypothetical flat space-time.

Wormholes as time machines

The biggest surprise in all this is that if the laws of physics do permit wormholes suitable for space travel, then they provide a simple means of time travel as well. A wormhole may be turned into a time machine by keeping one mouth of the hole fixed with respect to the distant stars, while moving the other. (From outside, a wormhole mouth is an ordinary gravitating body. You could move it using the gravitational attraction of another body, or by giving it an electric charge and moving it with electric fields.) Clocks fixed to the moving mouth advance more slowly than those at the stationary mouth; they undergo "time dilation" with respect to distant clocks, a well-known effect predicted by the special theory of relativity. However, they remain linked to clocks at the stationary mouth through the wormhole.

Enter the wormhole at the moving mouth when clocks there read 12:00 and you will emerge from the stationary mouth with the clocks there reading just after 12:00. The accumulated time dilation of the moving mouth, then, makes backward time travel possible. Eventually, you can travel from the stationary mouth to the moving one, through the intervening space, and reach the moving mouth when clocks there read an earlier value than those at the stationary mouth did when you left. Travel back through the hole, and you arrive at your starting point at an earlier time than you left (see Figure 2). You have made a journey, through the wormhole, back in time.

This is not the unrestricted time-travel of science fiction. There is a surface in the wormhole space-time, defined, as shown in Figure 2, by the first cyclic journey when no time has elapsed, before which no backward time-travel can take place. But in the space-time beyond this boundary cyclic trips through time are possible. Hence we must face all the paradoxes associated with time-travel, or re-evaluate our familiar ideas of causality and time evolution.

Notions of causality—that causes precede effects, that the past determines the present and the future, and so on—are deeply ingrained in scientific thought. The team from Caltech, in collaboration with Igor Novikov of the Space Research Institute, Moscow, and other physicists, are examining the implications for these ideas of wormhole time machines. They supplement causality with the principle of consistency: the evolution of a physical system should be self-consistent, even when you include influences from the future acting back in time. This means that if you travel back in time and attempt to

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

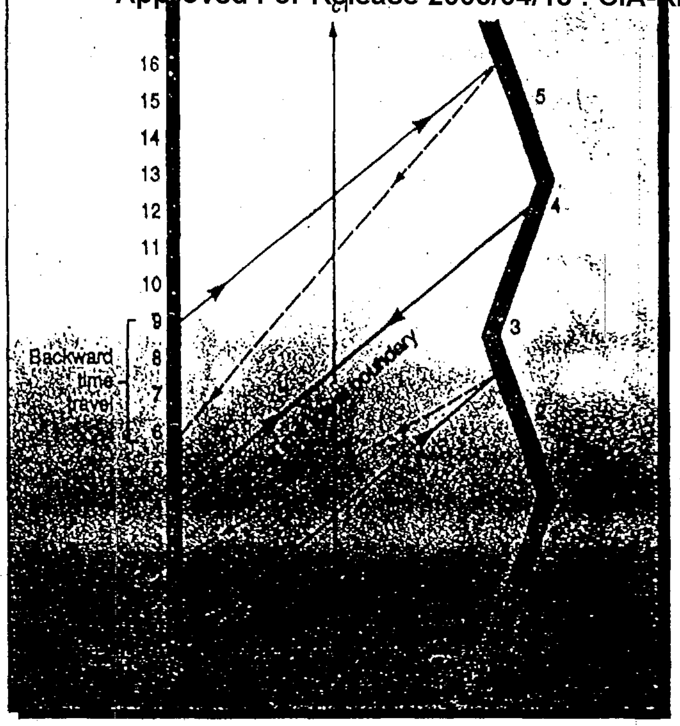


Figure 2 *A wormhole time machine. You travel from the stationary mouth S to the moving mouth M through the intervening space at speeds up to that of light* c. *You return immediately through the wormhole to the point where* t_s *equals* t_m *where you entered M. After the boundary shown, such journeys can take you back in time*

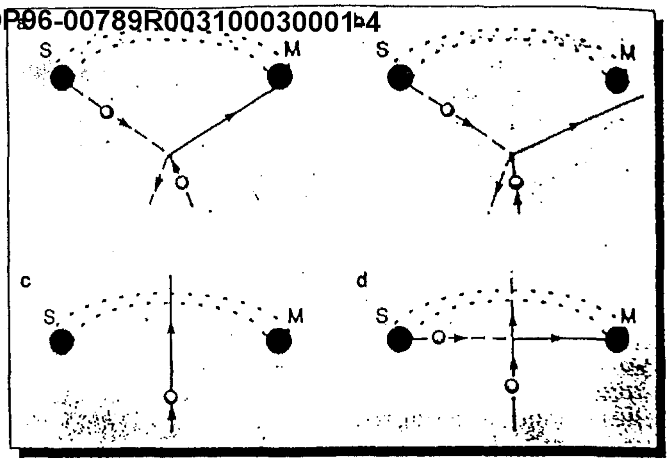


Figure 3 *Consistency with a wormhole time machine. For the right initial velocity as in* **a**, *the billiard ball emerges from the stationary mouth S, after backward time travel, just in time to knock itself back into the moving mouth M. Solution* b *is inconsistent: the billiard ball fails to enter M to produce the earlier collision. More than one consistent outcome is possible for the same initial conditions. In* c, *the ball passes between the wormhole mouths without interaction; in* d, *it emerges from S after backward time travel to knock itself into M*

shoot your parents before your birth, your gun misfires or you miss; the sequence of events already includes the effects of your attempt.

The researchers find that a simple "free" field in space-time with a wormhole evolves in a consistent and well-determined way from any initial conditions specified well before the wormhole's time-travel boundary. Obtaining consistent evolution from conditions specified after that boundary is more of a problem; the initial conditions may have to be restricted or specified at various times.

Interacting systems present further complications, as illustrated by the problem of colliding "billiard balls" in a wormhole space-time. Consistency restricts the range of possible initial conditions (see Figure 3 a and b). With the right initial velocity, a ball can be knocked into the wormhole and travel back in time to knock itself into the hole, in a consistent way. With the wrong velocity, the ball emerging from the time machine fails to knock itself into the hole; this is excluded by the principle of consistency. But more is needed. Some initial conditions imply more than one distinct, consistent solution, as in Figure 3 c and d. A ball can pass undisturbed between the wormhole mouths, or it can be deflected into the hole, travel back in time, and emerge to cause the deflection. Both are consistent outcomes of the same initial conditions. The researchers from Caltech and Moscow resolve this paradox by treating the balls according to the laws of quantum mechanics. In this light, both outcomes occur with some probability, but the probability distribution for the balls evolves in a unique and consistent way. While much more remains to be understood, it appears that time travel may be more physically reasonable than everyone had thought. ►

New space-time geometries for quantum gravity

THE DESCRIPTION of quantum gravity in which wormholes appear to play so large a role is a generalisation of the formulation devised by Richard Feynman in 1948 for ordinary quantum mechanics. This is called the "sum-over-histories" approach. For example, a classical particle travels from one position to another along a particular path, or "history". According to quantum mechanics, however, the particle has no definite path; the uncertainty principle forbids it. All that can be determined is the probability for the particle to go from one position to the other. This can be calculated by adding up contributions from all paths between the two positions.

For gravity, the interesting quantity, corresponding to the particle's position, is the geometry of space, three-dimensional space at some instant of time. A history is the evolution of the space geometry from one such instant to another, in other words, it is the four-dimensional geometry of space-time between the instants. In quantum gravity, a sum over such four-dimensional histories should give some measure of the probability for space to evolve from one three-dimensional geometry to another. Of course this is a rough, general description. Both the precise construction of such a sum and its interpretation are problems that have occupied theorists for decades, and still are not well understood.

Hawking and Coleman, among others, for a variety of technical reasons, do not construct their sum-over-histories from space-time geometries connecting two "instants of time". Instead, they use geometries with four dimensions which are all spatial, and the given three-dimensional geometries make up the boundaries of the four-dimensional one. Hawking claims this is the only way to obtain a sensible sum-over-histories for quantum gravity—though not everyone agrees with him—and has used this approach to evaluate the sum for very simple models of the Universe.

The theorems that forbid the creation of a wormhole in a smooth, classical space-time, as mentioned in the main text, do not apply to these four-dimensional spatial geometries. Thus wormholes branching off and rejoining the larger space contribute to the sum-over-histories in this approach to produce the effects claimed by Hawking and Coleman. Whether the same effects are found if the sum is formulated differently—using actual space-time geometries [illegible]—is uncertain. Some researchers contend that they are, others that they are not; the results depend sensitively on technical assumptions in the calculation.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

➤ One problem that remains unsolved is that of constructing a wormhole in the first place. Theory shows, for example, that you cannot create a wormhole in a smooth, classical space-time, with well-defined time directions everywhere, unless the space-time geometry already allows travel in time. Most theorists conjecture that on very small scales the geometry of space-time fluctuates in accord with the quantum uncertainty principle, giving rise to a "foam" of tiny wormholes. Perhaps a macroscopic wormhole could be obtained by enlarging one of these in some way. Only here does the matter of traversable wormholes and time-machines touch upon the physics of the other sort of wormhole—the microscopic wormhole.

In the late 1950s, John Wheeler, then at Princeton University, was already proposing that elementary particles might consist of microscopic wormholes threaded by electric or other field lines. This did not prove a useful description, but, since then, theorists have held that space-time should be permeated by wormholes on scales at which quantum mechanics affects gravity. This happens near the "Planck scale", around 10^{-35} metres. These wormholes should play an important role in any quantum theory of gravity.

In 1987, Stephen Hawking derived some concrete consequences of this; his results indicated that such wormholes modify quantum mechanics and alter the constants of nature in unpredictable ways. In 1988, Sidney Coleman of Harvard University contested Hawking's conclusions, though not his calculations, claiming instead that quantum wormholes actually fixed the values of certain physical constants in a dramatic fashion. Other theorists quickly joined in, some supporting Coleman's conclusions, some denying them.

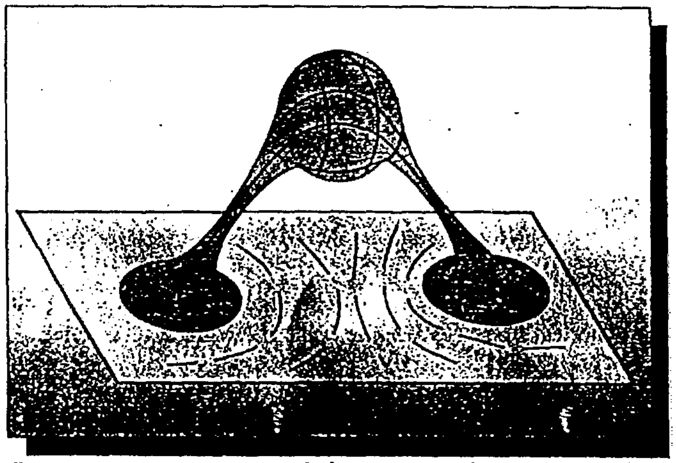


Figure 4 *A quantum wormhole geometry. The interior of the wormhole is a "baby universe" separated from a larger space*

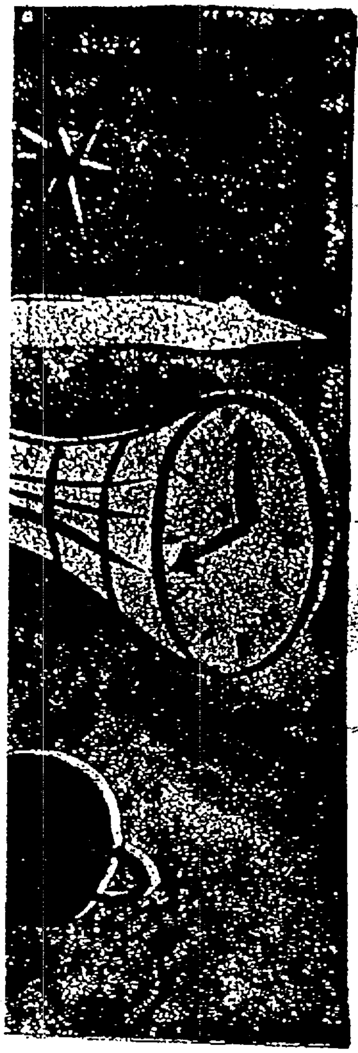
Hawking judged that a proper quantum-mechanical treatment of gravity (see Box on previous page) should include the effects of microscopic wormhole geometries such as the one shown in Figure 4. Such a geometry represents a "baby universe" (the interior of the wormhole) branching off and rejoining the larger Universe. He calculated that the contribution of these to physical processes was equivalent to a simple interaction between elementary particles and baby universes. He concluded that such an interaction would cause a loss of information (down the wormholes) from quantum systems in the larger Universe. The interaction would also shift the values of the constants in the original theory describing the elementary particles, by amounts that depended on the internal states of all possible baby universes. This means that even if we had a comprehensive "theory of everything", it would be useless. Nothing could be calculated from it without first making an infinite number of measurements to determine all the shifts caused by the baby universes.

Coleman's camp disputed this. They argued that Hawking's "loss of information" would not be observable. They went on to a much more startling conclusion: that the shifting of physical constants by baby universes could solve the long-standing "cosmological constant problem", and more.

The cosmological constant is the coefficient of a term in Einstein's gravitational field equations. It can be interpreted as an "energy density of the vacuum", a density that remains constant as the Universe expands or contracts. (Unlike matter or radiation, the vacuum does not become more or less dense

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

...or Release 2003/04/18 : CIA-RDP96-00789R003100030001-4



as the volume of the Universe [illegible] remains the vacuum.) Einstein put the constant into his equations in order to obtain a solution describing the Universe as it was believed to be prior to the late 1920s—filled with matter, but static. Einstein's motives for using the constant have since vanished—we know the Universe is expanding—and observations show it to be very small or zero, but theorists are still having trouble getting rid of it.

Elementary particle physics predicts vacuum energies arising from quantum fluctuations, like the Casimir energy mentioned earlier. The total energy density is typically 120 orders of magnitude larger than is consistent with the observations. To reconcile this, the theorists need to arrange for the contributions from different types of particles to cancel each other to 120 decimal places (unlikely) or to find some other way to get rid of the constant. Hawking suggested in 1984 that quantum gravity might do this; Coleman placed the idea on a firmer footing by invoking the effects of wormholes.

The wormholes contribute to what is called "sum-over-histories" in a quantum description of gravity. This gives the probability of a physical process in terms of a sum over all possible "paths" that the process can take. Coleman argued that if you take into account the contributions of microscopic wormholes in the "sum-over-histories" (see Box) for quantum gravity, it is completely dominated by histories in which the cosmological constant, in large regions of the Universe like our own, is zero. Any physical observation that we make to measure the constant must, therefore, give a zero result. Moreover, the completely dominant histories are also those for which Newton's gravitational constant, and other physical quantities appearing in the sum, take their minimum values. These requirements should determine all the internal states of the baby universes—all of Hawking's shifts in the physical constants—hence all the values of all those constants of nature. No wonder Coleman calls this "the big fix".

The possibility that quantum gravity could have such dramatic effects, and that they might be calculated, has drawn many theorists to the subject. It has become a hotbed of activity in the past two years. Many variations on the original calculations, and new calculations, have been published—to test the validity of the assumptions that were made, to examine in full detail particular models of wormholes, or to search for particular observable effects.

The arguments of Coleman and others have flaws, however, that may invalidate all their conclusions. Some careful calculations of the "sum-over-histories" that Coleman uses indicate that the histories with zero cosmological constant do not dominate as he claims, but are actually suppressed. It is not even certain that the whole approach to quantum gravity used by Coleman, Hawking and others is well defined. William Unruh of the University of British Columbia has found particularly devastating results along these lines. He claims that Coleman and Hawking omit a whole class of histories from their sum; when these are included the sum fails to give a finite prediction for any physical process. If Unruh is correct, then microscopic wormholes become a *reductio ad absurdum* for this approach. (That would be nearly as significant as solving the cosmological constant problem, but it is not a result most physicists would like.)

Even if Coleman's calculations are correct, the theory could still founder when compared with observations. If the theory forces Newton's constant to a minimum value that turns out to be zero or negative, it is undone. Recent results also suggest that the theory may predict masses for elementary particles in flagrant disagreement with their measured values, or a density of wormholes in space-time large enough to conflict with well-known particle physics.

The controversy is far from over. Microscopic wormholes may provide a breakthrough in our understanding of quantum gravity, or they may completely invalidate our present models, or they may yet prove to be a dead end.

No one has ever observed a wormhole, large or small. All the ideas that I have described are extensions of theory, reasonably well-founded in the classical case, but less so in the quantum case. It is the hope of every physicist working on either subject to come up with physical effects resulting from these speculations that will bring them within reach of the experimenter or the observer. □

Ian Redmount is a research associate in physics at Washington University, St Louis, Missouri. He was formerly a PhD student of Kip Thorne at Caltech.

WHERE ELSE COULD YOU BE SCIENTIFICALLY PARKED FOR £3?

Yes, for just £3.00 per sq.ft. you can rent a top-quality unit in Aberystwyth Science Park, close to the University College of Wales' superb academic facilities and in the middle of beautiful surroundings.

Trying to do the same thing in the South-East could cost you up to £20 per sq.ft.,with a business rates bill on top of that.

Naturally, our business rates are low too - giving you a possible saving of £85,000 per annum on a 5,000 sq.ft. unit!

And you'll make other gains too - countryside that's an inspiration in itself, and a ready source of enthusiastic, well qualified staff.

See for yourself why Aberystwyth Science Park offers growth-conscious scientific enterprises the ideal growth medium - but don't delay, or it'll be someone else who makes the £85,000 saving...

Please send me details of Aberystwyth Science Park.

I am interested in: 500 sq.ft. ☐ 2,000 sq.ft. ☐ 5,000 sq.ft. ☐ units

NAME ____________ POSITION ____________

COMPANY ____________

ADDRESS ____________

____________ POSTCODE ____________

TELEPHONE ____________

Send to: Dept A1, 6 Aberystwyth Science Park, Aberystwyth, Mid Wales SY23 3AH.
Or telephone us on 0970 615779.

Rural Wales
The New Country

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The Zeno's paradox in quantum theory

B. Misra and E. C. G. Sudarshan*

Center for Particle Theory, University of Texas at Austin, Austin, Texas 78712

(Received 24 February 1976)

We seek a quantum-theoretic expression for the probability that an unstable particle prepared initially in a well defined state ρ will be found to decay *sometime during a given interval*. It is argued that probabilities like this which pertain to continuous monitoring possess operational meaning. A simple natural approach to this problem leads to the conclusion that an unstable particle which is continuously observed to see whether it decays will never be found to decay! Since recording the track of an unstable particle (which can be distinguished from its decay products) *approximately* realizes such continuous observations, the above conclusion seems to pose a paradox which we call Zeno's paradox in quantum theory. The relation of this result to that of some previous works and its implications and possible resolutions are briefly discussed. The mathematical transcription of the above-mentioned conclusion is a structure theorem concerning semigroups. Although special cases of this theorem are known, the general formulation and the proof given here are believed to be new. We also note that the known "no-go" theorem concerning the semigroup law for the reduced evolution of any physical system (including decaying systems) is subsumed under our theorem as a direct corollary.

1. INTRODUCTION

The object of this paper is to discuss a seemingly paradoxical result in quantum theory concerning temporal evolution of a dynamical system under continuous observation during a period of time. For reasons that will become clear shortly we call this complex of deductions *Zeno's paradox in quantum theory.*

Let us consider schematically the theory of an unstable quantum system. Naturally the states corresponding to the decay products also should be included in the space of all states which we take to be a Hilbert space $\mathcal{H}$. Let us denote the (orthogonal) projection onto the subspace spanned by the *undecayed* (unstable) states of the system by E. This projection E thus represents the observable that corresponds to the "yes-no experiment" for determining whether the system is in an *undecayed* state or in a *decayed* state. The evolution in time of the states of the total system will be described by a unitary group $U(t) \equiv \exp(-iHt)$ labeled by the real time parameter t. In this setting the quantity

$$q(t) = \mathrm{Tr}[\rho U^*(t) E U(t)] \tag{1}$$

is interpreted as the probability, *at the instant t*, for finding the system undecayed when at time 0 it was prepared in the state ρ. Correspondingly, the probability that, *at the instant t*, the system will be found to have decayed is the complementary quantity

$$p(t) = 1 - q(t) = \mathrm{Tr}[\rho U^*(t) E^{\perp} U(t)], \quad E^{\perp} = I - E. \tag{2}$$

All these are, of course, standard.

Quantum theory, in fact, provides an unambiguous algorithm for computing the probability distributions of *time*, given the knowledge of the initial state of the of) *time*, given the knowledge of the initial state of the same system and its law of time evolution. Expressions (1) and (2) are only particular instances of this well-known algorithm.

In contrast to the above-mentioned probabilities which refer to a specified instant of time we may consider the following probabilities for which quantum theory has no ready expressions:

(1) The probability that the system prepared in the undecayed state ρ at time 0 is found to decay *sometime during the interval* $\Delta = [0, t]$. We denote this by $P(0, t; \rho$

(2) The probability $Q(0, t; \rho)$ that no decay is found *throughout the interval* Δ when the initial state of the system was known to be ρ.

(3) The probability that the system prepared initially in the state ρ will be found to be undecayed *throughout* $[0, t_1] \equiv \Delta$, but found to decay *sometime* during the subsequent period $[t_1, t] \equiv \Delta_2$, $0 < t_1 < t$. We denote this by $R(0, t_1, t; \rho)$.

It is important to distinguish the probabilities $Q(0, t; \rho)$ from $q(t)$ since there is the temptation to identify them [and hence also $P(0, t; \rho)$ with $p(t)$].[1] The probability $q(t)$, however, refers to outcomes of measurement of E at the time t, *the system being left unobserv* after the initial state preparation until t.

The operational meaning (if any!) of the probabilities P, Q, R on the other hand is to be found in terms of the outcomes of continuously ongoing measurement of E during the entire interval of time Δ. The notion of such continuously ongoing observations (or, equivalently, measurements) is obviously an idealization.

We may consider the process of continuing observation as the limiting case of successions of (practically) instantaneous measurements (of E) as the intervals between successive measurements approach zero. Since there does not seem to be any principle, internal to quantum theory, that forbids the duration of a single measurement or the dead time between successive measurements from being arbitrarily small, the process of continuous observation seems to be an admissible process in quantum theory.

It may be argued, however, that what are measurabl or not are governed not only by the fundamental principles of quantum theory but also by the actual constituen of the real world and their interactions.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The concept of continuous observation would indeed be bereft of any physical meaning *if* it could be established that the fundamental constituents of the real world and the interaction between them are such as to exclude the possibility of arbitrarily frequent observations. But, on the one hand, we cannot claim as final our present knowledge of the constituents and interactions of the real world. On the other hand, to agree that there is a limitation on the frequency of observation amounts to claiming the existence of an elementary and indivisible unit of time. Though the existence of an elementary interval of time is an exciting possibility, it is not part of the currently accepted and tested physical theories.

We, thus, feel that the notion of continuous observation should be accepted, at least for the present, as physically meaningful and quantum theory should be pressed to yield an answer to questions relating the probabilities pertaining to such observations.

Continuous observation processes seem to be realized in practice also, at least approximately, by the tracks of unstable charged particles in bubble chambers and other detecting media. The observation of the track amounts practically to a more or less continuous monitoring of the existence of the unstable particle and thus a measurement of E during the period of the particle's flight through the detection chamber. We are therefore led to accept as operationally meaningful the $P(0,t;\rho)$, $Q(0,t;\rho)$, and $R(0,t_1,t;\rho)$. To be a *complete* theory, quantum theory must provide an algorithm for computing these probabilities.

In the next section we describe what appears to be the natural approach to determining quantum-theoretic expressions for these probabilities. Our investigation leads to the paradoxical result mentioned at the beginning of this section: An unstable particle observed continuously whether it has decayed or not will *never* be found to decay! Since this evokes the famous paradox of Zeno denying the possibility of motion to a flying arrow, we call this result the Zeno's paradox in quantum theory

In fact, if E is taken to be the projection to the set of localized states of a particle (or, a quantum arrow) in a given region D of space, then one concludes that the particle will never be found to arrive in a disjoint region D' provided it is *continuously observed* whether it has entered D' or not: The "arrow" cannot move to where it is not!

This result acquires an even more picturesque and paradoxical formulation when it is applied to the "hellish contraption" considered in the Schrödinger's cat paradox.[2] It may be recalled that the contraption consists of an unstable (quantum) particle placed in a box equipped with an efficient counter and a cat inside a steel chamber. If the particle decays, the counter triggers and, in its turn, activates a tiny hammer which breaks a container of cyanide in the steel chamber. Monitoring the vital functions of the cat amounts to observing if the particle has decayed or not. In view of the Zeno's paradox formulated above, should we conclude that the particle will never decay? Will the cat escape the cruel death awaiting it, against which it has no defense, provided its vital signs are constantly watched with loving care?

The mathematical transcription of Zeno's paradox is a structure theorem concerning a class of strongly continuous semigroups. This theorem is formulated and proved in Sec. 3 of this paper and may possess some intrinsic interest apart from its application in the present context of a theory of continuous observation. As a by-product of this investigation we find that the known result[3] concerning the incompatibility of the semiboundedness of the Hamiltonian H with the requirement that $E\exp(iHt)E$ form a strictly contractive semigroup can be subsumed under the above-mentioned theorem as one of its direct corollaries.

Some of the implications and possible resolutions of the quantum Zeno's paradox will be briefly discussed in the concluding section of the paper.

Finally it may be mentioned that the conclusion called here the Zeno's paradox in quantum theory has been noted in some previous works,[4-6] but the present analysis of the problem is carried out in a more general and mathematically rigorous setting than the previous works. This, we feel, is not merely a dispensable luxury, but is necessary to locate the precise assumptions on which the "Zeno's paradox" rests.

2. QUANTUM THEORETICAL EXPRESSIONS FOR $P(0,t;\rho)$ AND RELATED PROBABILITIES

The three probability functions P, Q, R introduced in the previous section relate to the results of continuous observation throughout an interval of time. By their very definitions they must obey the relations

$$P(0,t;\rho)+Q(0,t;\rho)=1$$

and

$$R(0,t_1,t;\rho)=Q(0,t_1;\rho)\,P(0,t-t_1;\rho_1),$$

where ρ_1 is the state in which the system (prepared initially in the state ρ) finds itself at t_1 after being continuously observed *and* found to be undecayed throughout $[0,t_1]$. We may therefore concentrate our attention on calculating Q and ρ_1.

We start with the system in the state ρ and make a series of $n+1$ measurements, which are idealized to be instantaneous, at times 0, t/n, $2t/n,\ldots,(n-1)t/n$, and t. We seek the probability $Q(\Delta,n;\rho)$ that it be found undecayed in *each* of these measurements. It is natural to assume that $Q(\Delta;\rho)\equiv Q(0,t;\rho)$ can be evaluated as the limit of $Q(\Delta,n;\rho)$ when $n\to\infty$, provided the limit exists.

Let us denote by $\rho(n,t)$ the state in which the system finds itself after the $(n+1)$ measurements at 0, t/n, $2t/n,\ldots,t$ *and* being found to be undecayed in each of these measurements. Now, according to the orthodox theory of measurement, if a measurement of E on the system is carried out yielding the result "yes" (that is, "undecayed"), then the state of the system collapses to a new (unnormalized!) state ρ' of the form

$$\rho'=\sum_j A_j^*\rho A_j \tag{3}$$

with

$$\sum_j A_j^* A_j=E. \tag{4}$$

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

The collapsed state ρ' given by (3) is, in general, not uniquely determined by the measured observable E and the observed outcome but depends also on the details of the measuring apparatus. This circumstance is reflected in the nonuniqueness of the operators A_j satisfying (4).

The mapping (3) of the density matrices is very closely related to the "completely positive maps" defined by

$$\rho \to \sum_\alpha V_\alpha \rho V_\alpha^* = \mathcal{A}(V_\alpha)\rho,$$

$$\sum_\alpha V_\alpha^* V_\alpha = 1.$$

The "state collapse" caused by "nonselective" measurements of E is described by such maps. They will be considered in a future publication in the context of repeated and continuous nonselective measurements.

Quantum theory envisages also the possibility of *ideal* measurements under which the collapse of the state proceeds according to the simple law

$$\rho \to \rho' = E\rho E \qquad (5)$$

when the measurement of E on the state ρ yields the result "undecayed."

The considerations of this paper will be restricted to such ideal measurements only, since in such cases we can exploit the positive definiteness of the Hamiltonian in a direct manner. If we were to consider the more general collapses (3) we would have to proceed more indirectly using the von Neumann–Liouville generator which is however not positive definite. The study of the probabilities $Q(\Delta, \rho)$, etc. would then involve new technical problems obscuring the essentials of Zeno's paradox. We plan to present the study of the more general situation in a subsequent paper.[7]

Accordingly, to determine $\rho(n, t)$ we allow the system to collapse at each measurement according to (5) but at the intervening time intervals it undergoes the usual Schrödinger time development. The (unnormalized) state $\rho(n, t)$ is then easily seen to be

$$\rho(n, t) = T_n(t)\rho T_n^*(t), \qquad (6)$$

where

$$T_n(t) = [EU(t/n)E]^n \equiv [E\exp(-iHt/n)E]^n. \qquad (7)$$

Moreover, it is also easy to show that the standard interpretation of the quantum theoretical formalism entails the formula

$$Q(n, \Delta; \rho) = \mathrm{Tr}[T_n(t)\rho T_n^*(t)]. \qquad (8)$$

In fact, (8) is a special case of a more general formula for the probability connections between several successive observations.[8] It is important, however, to bear in mind that the general formula discussed in[8] [and, *a fortiori*, formula (8)] holds only under the assumption that the successive measurements under consideration are *ideal* in the sense described above. For nonideal successive measurements these formulas do not yield correct probability connections.

Returning to ideal measurements we have to proceed to the limit for $n \to \infty$. We define

$$\rho(t) = \underset{n\to\infty}{\text{s-lim}}\, \rho(n, t), \qquad (9)$$

$$Q(\Delta; \rho) = \lim_{n\to\infty} Q(\Delta, n; \rho), \qquad (10)$$

provided the limits on the right-hand side exist. Hence, if the limit

$$\underset{n\to\infty}{\text{s-lim}}\, T_n(t) \equiv \underset{n\to\infty}{\text{s-lim}}\, [EU(t/n)E]^n = T(t)$$

exists for $t \geq 0$, then we may make the identification

$$\rho(t) = \{\mathrm{Tr}[T(t)\rho T^*(t)]\}^{-1} \cdot T(t)\rho T^*(t) \qquad (11)$$

for the resultant (normalized) state obtained as a result of *continuous observation* and verification that the system remained *undecayed throughout the interval*. The probability $Q(\Delta; \rho)$ for this outcome is given by

$$Q(\Delta; \rho) = \lim_{n\to\infty} \mathrm{Tr}[T_n(t)\rho T_n^*(t)] = \mathrm{Tr}[\rho T^*(t)T(t)]. \qquad (12)$$

Once $Q(\Delta; \rho)$ is obtained in this manner we may calculate $P(\Delta; \rho)$ to be

$$P(\Delta; \rho) = \mathrm{Tr}[\rho(I - T^*(t))T(t)]. \qquad (13)$$

For a given group $U(t)$ of time-evolution the existence of the operator $T(t)$ for $t \geq 0$ imposes a nontrivial restriction on the projection E. This restriction may be viewed as a *necessary* condition in order that the observable represented by E admits a *continuous ideal* measurement.

It is known[9] that the operators $T(t)$ (if they exist) form a strongly continuous semigroup for $t > 0$. The continuity of $T(t)$ at $t = 0$ does not generally follow from the existence of $T(t)$, but on physical grounds, we shall *assume it*;

$$\underset{t\to 0_+}{\text{s-lim}}\, T(t) = E. \qquad (14)$$

This condition expresses the essentially desirable requirement that the probability $Q(\Delta; \rho)$ given by (12) approaches the probability $\mathrm{Tr}(\rho E)$ as $t \to 0_+$ that the system is undecayed initially.

To prove the existence of $T(t)$ and its continuity at the origin, (14), in specific examples of physical interest poses nontrivial mathematical problems. We hope to consider these in a subsequent paper.

3. ZENO'S PARADOX IN QUANTUM THEORY

In the preceding section we arrived at formula (13) for the probability $P(\Delta; \rho)$ that the system prepared initially in the undecayed state ρ will be observed to decay *sometime* during the interval $\Delta = [0, t]$. Despite the natural derivation of (13) we now show that the probability $P(\Delta; \rho)$ vanishes for all finite intervals Δ provided that the initial state was undecayed,

$$\mathrm{Tr}(\rho E) = 1. \qquad (15)$$

We are thus led to the paradoxical conclusion that an unstable particle will not decay as long as it is kept under [illegible] ...er it decays

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

or not. The mathematical transcription of this statement is the following theorem.

Theorem 1: Let $U(t) \equiv \exp(-iHt)$, t real, designate a strongly continuous one-parameter group of unitary operators in the (separable) Hilbert space $\mathcal{H}$. Let E denote an orthogonal projection in $\mathcal{H}$. Assume that:

(i) The self-adjoint generator H of the group $U(t)$ is *semibounded*.

(ii) There exists an (antiunitary) operator θ such that

$$\theta E \theta^{-1} = E,$$

$$\theta U(t)\,\theta^{-1} = U(-t) \quad \text{for all } t.$$

(iii) $\text{s-lim}_{n\to\infty} [EU(t/n)E]^n \equiv T(t)$ exists for all $t \geq 0$.

(iv) $\text{s-lim}_{t\to 0_+} T(t) = E$.

Then $\text{s-lim}_{n\to\infty}[EU(t/n)E]^n \equiv T(t)$ exists for *all* real t and possesses the following properties:

(a) The function $t \to T(t)$ is strongly continuous and for *all* real t and s satisfies the semigroup law:

$$T(t)\,T(s) = T(t+s),$$

(b) and

$$T^*(t) = T(-t).$$

Remarks: (1) The conclusions of the theorem imply the relation:

$$T^*(t)\,T(t) = E \quad \text{for } \textit{all real } t \tag{16}$$

so that $P(\Delta;\rho) = \mathrm{Tr}[\rho(I-E)] = 0$ for all ρ satisfying (15).

(2) With θ interpreted as the time-reversal (or CPT) operation, the assumption (ii) of the theorem turns out to be only a weak version of T or CPT invariance of the theory. Moreover it should be noted that assumption (iii) is used only once in the proof for concluding the existence of the strong limit (iii) for $t < 0$ as well.

(3) It is easy to give a relatively elementary proof of the theorem under the additional assumption that E is a one-dimensional projection onto a vector in the domain of H. The theorem is also known to hold in the special case that $H = -\nabla^2$ and E is given by

$$(E\psi)(x) = \chi(x)\,\psi(x), \quad \psi \in L^2(R^3),$$

where χ is the characteristic function of a (suitably smooth) region of R^3.[10] Theorem 1 generalizes this result to arbitrary semibounded H and arbitrary projection E.

(4) The semiboundedness of H is necessary: Consider the following counterexample. Let $V(t)$ be the operator family

$$(V(t)\,\psi) = \psi(x-t), \quad \psi \in L^2(R).$$

Let E be defined by

$$(E\psi)(x) = \begin{cases} \psi(x) & x \leq 0, \\ 0 & x > 0. \end{cases}$$

It is then easy to verify that

$$E V(t)\,E V(s)\,E = E V(t+s)\,E$$

for all $t, s \geq 0$, so that

$$\text{s-lim}_{n\to\infty} [EV(t/n)E]^n = EV(t)E = T(t)$$

for all $t \geq 0$, but

$$T^*(t)\,T(t) = E V(t^*)\,E V(t)\,E$$
$$= V^*(t)\,E V(t) \neq E \quad \text{for all } t > 0.$$

Thus the conclusion of the theorem is violated, though the assumptions in its formulation except the semiboundedness of the self-adjoint generator $V(t)$ are met. [Strictly speaking, assumption (ii) about the existence of θ is also not satisfied, but it was necessary only to prove the existence of $T(-t)$ and $T(-t)$ is trivially verified to exist in the present example.]

We now turn to the

Proof of Theorem 1: The existence of

$$T(t) \equiv \text{s-lim}_{n\to\infty} [EU(t/n)\,E]^n \tag{17}$$

for *all* real t follows immediately from the assumed existence of $T(t)$ for *positive* t and assumption (ii). In fact for $t \geq 0$

$$\begin{aligned} T(-t) &\equiv \text{s-lim}_{n\to\infty} [EU(-t/n)\,E]^n \\ &= \text{s-lim}_{n\to\infty} \theta[EU(t/n)\,E]^n\,\theta^{-1} \\ &= \theta T(t)\,\theta^{-1}. \end{aligned} \tag{18}$$

To prove assertion (b) we observe that

$$[EU(-t/n)\,E]^n = \{[EU(t/n)\,E]^n\}^*$$
$$\to T^*(t) \quad \text{weakly as } n \to \infty.$$

On the other hand,

$$[EU(-t/n)\,E]^n \to T(-t) \quad \text{strongly as } n \to \infty.$$

The assertion (b) follows immediately.

It remains to prove (a), especially assertion (16). Let us make a slight notational change and write

$$W(t) \equiv T(-t) = \text{s-lim}_{n\to\infty} [E\exp(iHt/n)\,E]^n. \tag{19}$$

The statement (a) can be transcribed into a corresponding statement on $W(t)$. We shall also assume, without loss of generality, that H is self-adjoint and positive.

For convenience in exposition we shall break up the proof of statement (a) into the following three lemmas:

Lemma 1: Let H be a positive self-adjoint operator of the Hilbert space $\mathcal{H}$ and let $F_n(z)$ be the operator-valued function defined by

$$F_n(z) \equiv [E\exp(iHz/n)\,E]^n. \tag{20}$$

Then

(1) $F_n(z)$ is defined and continuous (in the strong operator topology) for all complex z with $\mathrm{Im}z \geq 0$ and it is *holomorphic* in the open upper half-plane $\mathrm{Im}z > 0$.

(2) The function $F_n(z)$ has the integral representation

$$F_n(z) = \frac{(z+i)^2}{2\pi i}\int_{-\infty}^{\infty} \frac{F_n(t)}{(t+i)^2(t-z)}\,dt, \quad \mathrm{Im}z > 0. \tag{21}$$

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

(3)

$$\frac{1}{2\pi i}\int_{-\infty}^{\infty}\frac{F_n(t)\,dt}{(t+i)^2(t-z)}=0,\quad \mathrm{Im}\,z<0. \tag{22}$$

Proof of Lemma 1: The assertion (1) follows from the positive self-adjointness of H and its proof is standard. To prove assertion (2) we start with Cauchy's integral formula for the function $F_n(z)/(z+i)^2$ which is holomorphic in the open upper half-plane,

$$\frac{F_n(z)}{(z+i)^2}=\frac{1}{2\pi i}\oint_C\frac{F_n(z')}{(z'+i)^2(z'-z)}\,dz',\quad \mathrm{Im}\,z>0,$$

where C is *any* simple *closed* rectifiable contour enclosing the point z and contained entirely in the open upper half-plane. A similar integral representation holds of course for the holomorphic function $F_n(z)$ itself. But with the choice we have made the integrand vanishes faster than $|z'|^{-1}$ as $|z'|\to\infty$. Hence if we choose the closed contour C to be the axis running from $-\infty+i\epsilon$ to $+\infty+i\epsilon$ and an infinite semicircle we could rewrite the contour integral as an open line integral

$$F_n(z)=\frac{(z+i)^2}{2\pi i}\int_{-\infty}^{\infty}\frac{F_n(t+i\epsilon)}{(t+i+i\epsilon)^2(t-z+i\epsilon)}, \tag{21'}$$

$\mathrm{Im}\,z>\epsilon>0.$

The (operator) norm of this integrand is dominated by the integrable function

$$(1+t^2)^{-1}\cdot(\mathrm{Im}\,z-\epsilon_0)^{-1}$$

for all ϵ with $0\leq\epsilon<\epsilon_0<\mathrm{Im}\,z$. Moreover,

$$\text{s-}\lim_{\epsilon\to 0_+}F_n(t+i\epsilon)=F_n(t).$$

Hence the conditions for the application of Lebesgue's dominated convergence theorem for operator-valued integrals[11] are met and (21′) goes over to the desired representation (21) in the limit $\epsilon\to 0_+$. The relation (22) is similarly obtained from the vanishing of the contour integral

$$\frac{1}{2\pi i}\oint_C\frac{F_n(z')}{(z'+i)^2(z'-z)}\,dz'\quad\text{for}\quad \mathrm{Im}\,z<0.$$

Lemma 2: With the same notation as in Lemma 1 let us assume that

$$W(t)\equiv\text{s-}\lim_{n\to\infty}F_n(t)\equiv\text{s-}\lim_{n\to\infty}[E\exp(iHt/n)E]^n$$

exists for all real t. Then:

(1) $W(z)\equiv\text{s-}\lim_{n\to\infty}F_n(z)$ exists for all z with $\mathrm{Im}\,z>0$.

(2) The function $W(z)$ is holomorphic in the *open* upper half-plane and satisfies the semigroup composition law

$$W(z_1)\,W(z_2)=W(z_1+z_2). \tag{23}$$

(3) There exists a nonnegative and self-adjoint operator B and a projection G such that

$$GB=BG=B,$$

and

$$W(z)=G\exp(iBz)\,G,\quad \mathrm{Im}\,z>0. \tag{24}$$

Proof of Lemma 2: To prove (1) we start with the representation (21) for $F_n(z)$. By assumption, $W(t)\equiv\text{s-}\lim_{n\to\infty}F_n(t)$ for all real t and

$$\left\|\frac{F_n(t)}{(t+i)^2(t-z)}\right\|\leq\frac{(1+t^2)^{-1}}{|\mathrm{Im}\,z|}\quad\text{for all } n.$$

We can therefore apply again the Lebesgue theorem on dominated convergence and conclude that

$$W(z)=\text{s-}\lim_{n\to\infty}F_n(z)$$

exists and has the representation

$$W(z)=\frac{(z+i)^2}{2\pi i}\int_{-\infty}^{\infty}\frac{W(t)}{(t+i)^2(t-z)}\,dt,\quad \mathrm{Im}\,z>0. \tag{25}$$

From the well known Vitali's theorem[11] we can conclude that $W(z)$ is holomorphic in the open upper half-plane.

To prove the semigroup property of $W(z)$ we show first that this law holds for pure imaginary values,

$$W(is)\,W(it)=W(i(t+s))$$

for all positive t and s.

To this end, first consider the case where t and s are rationally related so that there exist positive integers p,q for which

$$\frac{s+t}{r(p+q)}=\frac{s}{rp}=\frac{t}{rq}$$

for all integers r. For such s,t we can deduce

$$\left[E\exp\left(-H\frac{t+s}{r(p+q)}\right)E\right]^{r(p+q)}=\left[E\exp\left(-H\frac{s}{rp}\right)E\right]^{rp}\left[E\exp\left(-H\frac{t}{rq}\right)E\right]^{rq}$$

which, in the limit $r\to\infty$ yields

$$W(i(t+s))=W(is)\,W(it).$$

Once this is established for rationally related positive s and t by continuity it can be extended to all positive s and t. Since $W(is)$ is holomorphic it is, *a fortiori*, continuous for $s>0$.

To prove assertion (3) we observe that the operators

$$W(is)\equiv\text{s-}\lim_{n\to\infty}[E\exp(-Hs/n)E]^n$$

form a semigroup of self-adjoint operators for $s>0$. According to a well known structure theorem for such semigroups[12] there exists a *self-adjoint nonnegative* operator B and a projection G such that

$$BG=GB=B$$

and

$$W(is)=G\exp(-Bs)\,G,\quad s>0.$$

The function $z\to G\exp(iBz)\,G$ is holomorphic in the open upper half-plane and assumes the same values $W(is)$ as the holomorphic function $W(z)$ for $z=is$ $(s>0)$. The uniqueness of holomorphic functions, then, immediately establishes the representation (24). The semigroup property (23) for z_1, z_2 in the open upper half-plane follows from (24) and the commutativity of G and B.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Lemma 3: Under the assumptions … we have the weak limit for operators along the real axis

$$\underset{\eta\to 0_+}{\text{w-lim}}\, W(s+i\eta) = W(s) \quad \text{for almost all real } s. \tag{26}$$

Proof of Lemma 3: To obtain this weak limit let us start from the integral representation (25) rewritten in the form

$$W(s+i\eta) = \frac{(s+i+i\eta)^2}{2\pi i}\int_{-\infty}^{\infty} \frac{W(t)}{(t+i)^2(t-s-i\eta)}\, dt, \quad \eta > 0.$$

On the other hand, from (22) and the Lesbegue dominated convergence theorem

$$0 = \frac{(s+i+i\eta)^2}{2\pi i}\int_{-\infty}^{\infty} \frac{W(t)}{(t+i)^2(t-s+i\eta)}\, dt, \quad \eta > 0.$$

Therefore,

$$W(s+i\eta) = \frac{(s+i+i\eta)^2}{\pi}\int_{-\infty}^{\infty} \frac{W(t)}{(t+i)^2}\, \frac{\eta}{(t-s)^2+\eta^2}\, dt.$$

For any two vectors ψ, ϕ in $\mathcal{H}$ we may write

$$(\psi, W(s+i\eta)\,\phi) = \frac{(s+i+i\eta)^2}{\pi}\int_{-\infty}^{\infty} \frac{(\psi, W(t)\,\phi)}{(t+i)^2}\, \frac{\eta}{(t-s)^2+\eta^2}\, dt.$$

Since the quantity $(\psi, W(t)\,\phi)/(t+i)^2$ considered as a function of t is integrable, it follows that

$$\lim_{\eta\to 0_+} (\psi, W(s+i\eta)\,\phi) = (\psi, W(s)\,\phi) \tag{27}$$

for almost all s.

To complete the assertion of Lemma 3 a technical difficulty is to be resolved. For a given pair ψ, ϕ of vectors, the assertion (29) has been shown to hold for almost all s. The exceptional set (of measure zero) where this result may not hold may appear to depend on the pair ψ, ϕ chosen. To show that there is at most a single *null* set outside which (27) holds for *all* pairs ψ, ϕ we proceed as follows: Let $\mathcal{D}$ be a *countable* dense subset of the separable Hilbert space $\mathcal{H}$ and let $\mathcal{N}$ be the union of the *countable* family of exceptional null sets corresponding to all pairs ψ, ϕ with $\psi \in \mathcal{D}$, $\phi \in \mathcal{D}$. This set $\mathcal{N}$ is a set of measure zero and the weak limit (26) holds everywhere outside this set for ψ, ϕ in $\mathcal{D}$, but then (27) will hold in the complement of $\mathcal{N}$ for all pairs ψ, ϕ not necessarily in $\mathcal{D}$. In fact, writing

$$A(s,\eta) = W(s+i\eta) - W(s)$$

we may obtain

$$(\psi, A(s,\eta)\,\phi) = (\psi - \psi_n, A(s,\eta)\,\phi) + (\psi_n, A(s,n)(\phi-\phi_n)) + (\psi_n, A(s,\eta)\,\phi_n).$$

We see that for s outside the exceptional set $\mathcal{N}$ the first two terms on the right-hand side tend to zero as $n\to\infty$, since we may choose

$$\underset{n\to\infty}{\text{s-lim}}\, \psi_n = \psi, \quad \underset{n\to\infty}{\text{s-lim}}\, \phi_n = \phi.$$

The third term, by hypothesis, goes to zero as $\eta \to 0_+$ since ψ_n, ϕ_n are chosen to lie in $\mathcal{D}$. [The proof of this lemma incorporates a suggestion due to K. Sinha.]

… (a) of the theorem may now be easily completed by combining the conclusions of the preceding lemmas,

$$W(s) = \underset{\eta\to 0_+}{\text{w-lim}}\, W(s+i\eta) = \underset{\eta\to 0_+}{\text{w-lim}}\, G\exp[iB(s+i\eta)]\,G$$
$$= G\exp(iBs)\,G \quad \text{for almost all real } s. \tag{24'}$$

Thus $W(s)\,W^*(s) = G$ for almost all real s. According to assumption (iv) in the statement of Theorem 1,

$$\underset{s\to 0_+}{\text{w-lim}}\, W(s)\,W^*(s) = \underset{s\to 0_+}{\text{w-lim}}\, T(-s)\,T^*(-s)$$
$$= \underset{s\to 0_+}{\text{w-lim}}\, T^*(s)\,T(s) = E. \tag{28}$$

Thus $G = E$ and we may rewrite (24) in the form

$$W(s) = E\exp(iBs)\,E \text{ for almost all } s, \qquad EB = BE = B. \tag{29}$$

But we can now strengthen this relation for $W(s)$ to read

$$W(s) = E\exp(iBs)\,E \quad \text{for } all\ s \tag{30}$$

in view of the strong continuity of $W(s)$. Combining (29) and (30) we immediately deduce the validity of assertion (a).

Although not of primary interest for the discussion in this paper, we recall the known result[3] that if H is semibounded, the operators $E\mathrm{U}(t)\,E$ cannot form a semigroup for $t \geq 0$ except in the event of E commuting with $\mathrm{U}(t)$ for all real t. We may subsume this result as a corollary to Theorem 1.

Corollary: Let the self-adjoint operator H be semibounded, let E be an orthogonal projection and let $\mathrm{U}(t)$ stand for $\exp(-iHt)$. If $\{E\mathrm{U}(t)\,E \mid t \geq 0\}$ form a semigroup, then

$$E\mathrm{U}(t) = \mathrm{U}(t)\,E \quad \text{for all real } t. \tag{31}$$

Proof: The semigroup property for $E\mathrm{U}(t)\,E$, i.e., the relation

$$E\mathrm{U}(t)\,E\mathrm{U}(s)\,E = E\mathrm{U}(t+s)\,E \quad \text{for } all\ real\ s, t \tag{32}$$

will imply

$$E\mathrm{U}(t)\,E\mathrm{U}^*(t)\,E = E \quad \text{for } all\ real\ t \tag{32'}$$

and hence

$$E \leq \mathrm{U}(t)\,E\mathrm{U}^*(t) \quad \text{for } all\ real\ t.$$

Multiplying this equation from the left by $\mathrm{U}(-t)$ and the right by $\mathrm{U}^*(-t)$ will yield

$$\mathrm{U}(-t)\,E\mathrm{U}^*(-t) \leq E \quad \text{for } all\ real\ t.$$

Together these two inequalities imply

$$E = \mathrm{U}(t)\,E\mathrm{U}^*(t) \quad \text{for } all\ real\ t$$

or, equivalently,

$$E\mathrm{U}(t) = \mathrm{U}(t)\,E \quad \text{for } all\ real\ t.$$

The proof of the corollary is thus reduced to the proof of (32) or (32').

Since the operators $E\mathrm{U}(t)\,E$ are assumed to form a semigroup for $t \geq 0$ and $[E\mathrm{U}(t)\,E]^* = E\mathrm{U}(-t)\,E$, for all positive integers n and *all real* t we have

$$[E\mathrm{U}(t/n)\,E]^n = E\mathrm{U}(t)\,E.$$

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Hence

$$\text{s-}\lim_{n\to\infty}[EU(t/n)E]^n = EU(t)E \equiv T(t)$$

exists for *all* real t and all the assumptions of Theorem 1 are verified, except for (ii). But as we have pointed out, this assumption itself was needed for the sole purpose of guaranteeing the existence of $T(t)$ for *all real* t. Thus we can safely conclude that Theorem 1 applies in this case also and hence (32) holds for *all* real t and s.

4. CONCLUDING DISCUSSION

What conclusions must we draw from Zeno's paradox in quantum theory? Is it a curious but innocent mathematical result or does it have something to say about the foundation of quantum theory? Does it, for example, urge us to have a principle in the formulation of quantum theory that forbids the continuous observation of an observable that is not a constant of motion?

The answer to the first two questions appears to depend on whether it is operationally meaningful to seek the probability that the particle makes a transition from a preassigned subspace of states $E\mathcal{H}$ to the orthogonal subspace $E^\perp\mathcal{H}$ *sometime* during a given period of time. We have endeavored to present arguments that such probabilities possess operational meaning in terms of the outcome of successive (in the limit, continuous) measurements of an appropriate quantum mechanical observable. If this is accepted, it follows that to be a *complete* theory quantum mechanics must provide an algorithm for computing these probabilities. The quantum Zeno's paradox shows that the seemingly natural approach to this problem discussed in the preceding sections leads to bizarre and physically unacceptable answers. We thus lack a trustworthy quantum-theoretic algorithm for computing such probabilities. Until such a trustworthy algorithm is developed the *completeness* of quantum theory must remain in doubt.

The lack of a trustworthy quantum-theoretic algorithm for probabilities like $P(0,t;\rho)$ is intimately connected with the difficulties involved in defining an operator of "arrival time" (or, more generally, "time of transition") in quantum theory.[4] Let us briefly discuss this problem in the context of the "time of decay" of an unstable particle.

From the definition of $P(0,t;\rho)$ it must have the following properties:

$$\begin{aligned}
&\text{(i) } P(0,t;\rho) \geq 0 \text{ for all } t \geq 0\\
&\text{(ii) } P(0,t;\rho) \geq P(0,s;\rho) \text{ for } t \geq s\\
&\text{(iii) } P(0,t;\rho) \to 1 \text{ for } t \to \infty\\
&\text{(iv) } P(0,t;\rho) \to \mathrm{Tr}[\rho E^\perp] \text{ for } t \to 0_+.
\end{aligned} \tag{33}$$

In addition, $P(0,t;\rho)$ may be assumed to be continuous as a function of t. If we were to succeed in finding a formula

$$P(0,t;\rho) = \mathrm{Tr}[\rho B(0,t)], \tag{34}$$

then the operator $B(0,t)$ would have the following properties:

(i) $B(0,t) \geq 0$,

(ii) $B(0,t) \geq B(0,s)$, $t \geq s$,

(iii) $B(0,t) \to I$ (strongly), $t \to \infty$,

(iv) $B(0,t) \to E^\perp$ (strongly), $t \to 0_+$,

(v) $B(0,t)$ is a strongly continuous function of t.

The family $B'(0,t) \equiv EB(0,t)E$ restricted to the subspace $E\mathcal{H}$ of the (unstable) undecayed states will then form a "generalized resolution of the identity" (GRI). Unlike the more familiar (projection-valued) resolution of the identity, a GRI does not necessarily determine a densely defined operator, but under some mild additional assumptions (which we need not specify explicitly here) the GRI $B'(0,t)$ will determine a Hermitian (though not necessarily self-adjoint) operator τ so that

$$(\psi, \tau\psi) = \int t\, d(\psi, B'(0,t)\psi) \text{ for all } \psi \in \mathcal{D}(\tau). \tag{35}$$

The operator τ thus defined may then be interpreted as the operator of "time of decay."

Conversely, if there exists a positive Hermitian "time of decay" operator τ associated with the subspace $E\mathcal{H}$ of undecayed states and $B'(0,t)$ denotes a GRI associated with it, then through (33) we may define the probability $P(0,t;\rho)$ which may be interpreted as the probability that the system prepared initially in the (undecayed) state ρ will be found to decay *sometime* during the interval $[0,t]$.

Looked at from this point of view, the Zeno's paradox thus strengthens and sharpens the pessimistic conclusion of Allcock[4] and others concerning the possibility of introducing an observable of "arrival time" in quantum theory. We must emphasize that in our study here, the conclusion is *not* based on certain *a priori*, but questionable, assumptions about τ; such, for instance, as the assumption that τ be "canonically conjugate" to the Hamiltonian, or that τ be a *self-adjoint* operator in the Hilbert space. In the literature such requirements were implicitly or explicitly placed on τ.

We have so far supposed that it is operationally meaningful to ask about probabilities such as $P(0,t;\rho)$ and $Q(0,t;\rho)$. We have also taken the stance that the observed tracks of unstable particles in a bubble chamber or photographic emulsion is in contradiction with the conclusion we have called Zeno's paradox in quantum theory. It is, however, possible to adopt one of the following attitudes:

(1) Probabilities such as $P(0,t;\rho)$ have no operational meaning: There is a fundamental principle in quantum theory that denies the possibility of continuous observation.

Since so far no such principle has been derived from or incorporated into quantum theory, this is not a satisfactory way of resolving the paradox at the present time.

(2) Zeno's paradox is based on the assumption that the continuous measurements are ideal measurements. But measurements (or, observations) involved in the recording of the track of an unstable particle in a detecting medium are nonideal in the sense of (3).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

This is a tenable view and it would deny the validity of Theorem 1 as stated and proved in this paper. It has the somewhat unsettling side effect that $P(0, t; \rho)$ and hence the "observed lifetime" of an unstable particle is not a property of the particle (and its Hamiltonian) only, but depends on the details of the observation process. At the present time we have no indication that this is so.

(3) The record of the track of a particle is not a *continuous* observation that the particle has not decayed, but only a discrete sequence of such observations; while Zeno's paradox obtains only in the limit of continuous observations.

While this is tenable, the sufficiently repeated monitoring of the particle should again lengthen the lifetime. There is, however, no indication that the lifetime of a (charged) unstable particle (say, a muon) is appreciably increased in the process of its track formation through bubble chamber. To shed additional light on this question a quantitative investigation of the effect of repeated monitoring on the lifetime of particles (in specific models) is in progress.[13]

(4) Natural though it seems, it is wrong to assume that the temporal evolution of a quantum system under continuing observation can be described by a linear operator of time-evolution such as $T(t)$. It can be described only in terms of a persistent interaction between the quantum system and the *classical measuring apparatus*. When this is done the quantum Zeno's paradox will either disappear or if it survives, at least, it will be understandable as the drastic change in the behavior of the quantum system caused by its continuous interaction with a classical measuring apparatus.

This point of view is at present only a program since there is no standard and detailed theory for the actual coupling between quantum systems with *classical measuring apparatus*. A beginning in this direction is made in a forthcoming paper.[14]

Having been forced into such unusual points of view by the quantum Zeno's paradox one is prompted to draw also some parallels between it and certain empirical findings in the study of human awareness. We shall present such close parallels between the quantum Zeno's paradox and the findings of sensory deprivation and other experiments pertaining to the study of consciousness in a separate publication.

In conclusion, it seems to us that the problems posed by Zeno's paradox have no clean cut resolution at the present time and deserve further discussions. It may also be reemphasized that the probabilities such as $Q(\Delta;\eta,\rho)$ [or $Q(\Delta;\rho)$] that pertain to the outcomes of successive measurements (or continuous measurements) depend on the law according to which "state collapses" occur at the time of measurements. Thus one may say that the "collapse of state vector" caused by measurement, which has haunted the foundation of quantum mechanics like an invisible ghost becomes visible through probabilities such as $Q(\Delta;\eta,\rho)$, etc. The probabilities pertaining to the outcomes of several successive (as well as continuous) measurements therefore deserve further theoretical as well as experimental study than they have received so far.

ACKNOWLEDGMENTS

It is a pleasure to thank Professors C. Chiu, M. Guenin, K. Gustafson, Ph. Martin, C Piron, and K. Sinha for their critical comments as well as constructive suggestions. We are also grateful to the referee for some useful comments, in particular, for bringing to our notice Refs. 5 and 6. One of us (B. M.) was partially supported by Swiss NSF during summer 1975.

*Work supported in part by Energy Research and Development Administration Contract E(40-1)3992.

[1]Another possible suggestion is to interpret

$$\frac{1}{t}\int_0^t \mathrm{Tr}[\rho U^*(s) E U(s)]\, ds$$

as the desired probability $Q(0,t;\rho)$. Apart from the fact that there is really no convincing reason for this interpretation this expression is *not* generally a monotone (decreasing) function of time t, a property which $Q(0,t;\rho)$ should possess.

[2]E. Schrödinger, Naturwissenschaften 23, 807 (1935).
[3]D. Williams, Commun. Math. Phys. 21, 314 (1971); L. Horwitz, J.P. Marchand, and J. Lavita, Rocky Mount. J. Math. 1, 225 (1975).
[4]G.R. Allcock, Ann. Phys. (N.Y.) 53, 251–348 (1969).
[5]W. Yorgrau in *Problems in Philosophy in Science*, edited by I. Lakatos and A. Musgrave (North-Holland, Amsterdam, 1968), pp. 191–92.
[6]H. Ekstein and A. Seigert, Ann. Phys. (N.Y.) 68, 509 (1971).
[7]B. Misra and E.C.G. Sudarshan, in preparation.
[8]E.P. Wigner, in *Foundations of Quantum Mechanics*, edited by B. d'Espgnat (Academic, New York, 1971), especially formulas 14 and 14(a), p. 16.
[9]P.R. Chernoff, Mem. Am. Math. Soc. 140 (1974).
[10]C.N. Friedman, Indiana Univ. Math. J. 21, 1001–11 (1972).
[11]E. Hille and R. Phillips, *Functional Analysis and Semigroups* (Am. Math. Soc. Colloq. Publ., Providence, R.I., 1957).
[12]Ref. 11, Sec. 22.3.
[13]C. Chiu, B. Misra, and E.C.G. Sudarshan, in preparation.
[14]E.C.G. Sudarshan and B. Misra, in preparation.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

PHYSICAL REVIEW D VOLUME 21, NUMBER 8 15 APRIL 1980

Meaning of an individual "Feynman path"

Y. Aharonov
*Department of Physics and Astronomy, University of South Carolina, Columbia, South Carolina 29208
and Department of Physics and Astronomy, University of Tel-Aviv, Ramat Aviv, Israel*

M. Vardi
Department of Physics and Astronomy, University of South Carolina, Columbia, South Carolina 29208

(Received 24 September 1979; revised manuscript received 6 February 1980)

In this article we give an operational meaning to an individual "Feynman path." In other words, we describe a process of dense measurements, made in temporal sequence, which check whether the particle moves along any given trajectory in space-time. We show that in this process the two assumptions of the space-time formulation of quantum mechanics, are realized: (a) The weight that the particle moves along a trajectory that has been checked by this process is the same for all trajectories, and in fact, we show that the particle follows, with probability 1, the trajectory that is being checked. (b) A phase is systematically accumulated, so that, at the end of this process, the state is multiplied by the familiar factor $\exp[(i/\hbar)\int L\,dt]$. As an immediate extension of the above formalism, we suggest a setup that measures the relative phase between any two trajectories. Finally, our approach points toward the possibility of extending the Feynman formalism in order to cover more general Hamiltonians.

INTRODUCTION

In 1948 Feynman published what is essentially a third formulation of quantum mechanics.[1] As is well known, the main idea in this formulation is to associate a probability amplitude, $\exp[(i/\hbar)\int L\,dt]$, with each possible classical trajectory that connects two space-time points [L is the classical Lagrangian, and the integral is evaluated along the path $X(t)$]. Each possible trajectory is assigned the same weight, and the sum (integral) over the contributions from all possible trajectories has to be carried out in order to get the transition amplitude between these two space-time points.

The concept of a trajectory in quantum mechanics is not a straightforward one because of the uncertainty principle involved.[2] Therefore, attempts have been made to apply the notion of continual observation[3] (which was mentioned by Feynman[1], p. 370) in order to investigate the operational meaning of the trajectories, which are the building blocks of this formalism.

This notion of continuous observations (or measurements) in quantum mechanics has recently attracted some attention,[3-7] because of the interesting features that were revealed. In particular, the following paradoxical property of such measurements was found: Consider the case where repeated observations are carried out in order to find the exact moment at which a transition from some initial state takes place. It turns out that, because of these observations, the transition never occurs. A particular example is the decay of an unstable system[4]: If the system is continuously observed, then it will never decay. Another example is the one of continuously observing a system that is initially confined to a finite space region,[5] and because of these observations it remains confined there. In this paper, we first show that the above paradoxical situation is a special case of a more general property of continuous measurements. Namely, if one checks by continuous observations if a given quantum system evolves from some initial state, to some other final state, along a specific trajectory in Hilbert space, the result is always positive, whether or not the system would have done so on its own accord.

When the above result is applied to the evolution of a state along a trajectory considered by Feynman, we find that the particle follows, with certainty, the trajectory that is being checked. Therefore, it is now meaningful to consider measurements of individual trajectories in space-time and their properties. In particular, the phase associated with the probability amplitude for motion along a given trajectory can be evaluated.

When this calculation is carried out, the phase turns out to be the one assumed by Feynman, (within a constant independent of the trajectory).

The possibility is therefore open to consider setups that measure directly the relative phase of any two individual trajectories (which have common end points); we describe in detail such a setup.

Finally, we point out that our analysis can be applied to more general trajectories in Hilbert space than those corresponding to the classical trajectories. We discuss briefly the relevance of this to the question of extending the Feynman formalism in order to cover arbitrary Hamiltonians.

 © 1980 The American Physical Society

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

CONTINUOUS MEASUREMENTS ON A SPIN-HALF SYSTEM

Consider a spin-half particle placed in a constant magnetic field pointing in the z direction with the initial direction of the spin in the $+x$ direction. The time evolution in this case is simply a rotation in the xy plane with the Larmor frequency (defined by $H=\vec{\mu}\cdot\vec{B}\equiv\frac{1}{2}\hbar\omega\hat{\sigma}_z$). If we now want to check when the spin moves out of its initial orientation by performing a dense set of measurements (in time) of $\hat{\sigma}_x$, we find that it does not move out at all (in analogy to the results discussed in the literature[3-7]). This can be seen as follows: In the infinitesimal time δt the free Hamiltonian rotates the direction of the spin, or equivalently changes the state by

$$|\sigma(\delta t)\rangle=\exp\left(-\frac{i}{\hbar}\hat{H}\delta t\right)|\sigma_x=+1\rangle$$

$$=\exp\left(-\frac{i}{2}\omega\hat{\sigma}_z\delta t\right)|\sigma_x=+1\rangle$$

$$=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ e^{-i\omega\delta t}\end{pmatrix},$$

where

$|\sigma(\delta t)\rangle$ is the eigenstate of the operator,

$\hat{\sigma}(\delta t)=\hat{\sigma}_x(0)\cos\omega\delta t+\hat{\sigma}_y(0)\sin\omega\delta t$,

with the eigenvalue plus one.

If we now measure $\hat{\sigma}_x$, the probability that the state collapses to $|\sigma_x=+1\rangle$ is

$$P_r(\sigma_x=+1)=|\langle\sigma_x=+1|\sigma(\delta t)\rangle|^2$$

$$=\cos^2\frac{\omega\delta t}{2}\simeq 1-\frac{(\omega\delta t)^2}{4}\quad(\text{for }\omega\delta t\ll 1).$$

If we repeat the same measurement at intervals of δt the probability that all of them will give the same result is obviously $[1-(\omega\delta t)^2/4]^N$. We now note that if $\delta t=T/N$, where T is the total period of observation, and we approach the limit of very dense measurements ($N\to\infty$), we end up freezing the state in its initial value $|\sigma_x=+1\rangle$, since

$$\lim_{N\to\infty}(1-1/N^2)^N=\lim_{N\to\infty}\exp(-1/N)=1\,.$$

For future reference we refer to this as *case a*.

Case b. Still with the same system as before, we show how it is possible to bypass the feature of freezing while insisting on continuous observations. To achieve this we use the so-called "deterministic observations,"[8] namely, we measure the dense set of operators defined as follows:

$$\hat{\sigma}_n=\hat{\sigma}_x\cos\alpha_n+\hat{\sigma}_y\sin\alpha_n\,,$$

where $\alpha_n=\omega n\delta t$, ω is the Larmor frequency, and $n=1,\ldots,N$. We then obviously find $\sigma_n=1$ for $n=1,\ldots,N$. Thus it seems that we have found a way to monitor the time evolution of a system without freezing it in its initial state.[9]

Case c. In this case we show that the seemingly innocent deterministic observations described above have unsuspected features. Consider again the spin-half particle, initially with $\sigma_x=+1$, but *without a magnetic field*. If we measure the same dense set of operators as in *case b*, we find the following:

The conditional probability to get $\sigma_n=1$ if $\sigma_{n-1}=+1$ is

$$P_r(n)=|\langle\sigma_n=+1|\sigma_{n-1}=+1\rangle|^2$$

$$=\frac{1+\cos\omega\delta t}{2}=\cos^2(\omega\delta t/2),$$

where we have used the eigenvector of $\hat{\sigma}_n$ (that belongs to $\sigma_n=+1$), namely

$$|\sigma_n=+1\rangle=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ e^{i\alpha_n}\end{pmatrix}.$$

Thus, for sufficiently large N, the probability of finding $\sigma_n=+1$ in all the measurements is essentially one. This is so even though no magnetic field is present; therefore, the only reasons for the spin rotation are those measurements. This result is quite surprising because of the accepted assumption that if the outcome of a measurement of some dynamical variable is certain (i.e., with probability one), then the state of the system was not disturbed.

The above conclusion can also be derived by the following simple argument: If we analyze *case a* from a rotating-frame-of-reference point of view so that in the new frame the inertial field exactly cancels the original constant magnetic field,[10] we end up in the same situation as in *case c* (in which the continuous measurements were made at the rate dictated by the Larmor frequency). Therefore, if we accepted that the measurements in *case a* froze the system in the state $|\sigma_x=+1\rangle$, we end up with the result of *case c* as a necessary consequence.

THE PARTICLE FOLLOWS A TRAJECTORY IN PHASE SPACE

We have shown in a simple example that it is possible to define a set of operators that represent a dense sequence of measurements, so that the initial state of the system evolves along the eigenstates of these operators, and the probability to get each of the corresponding eigenvalues is one. This idea can be generalized to more complicated systems, and we now apply it to obtain a

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

operational definition of a "Feynman path." We recall that in Feynman's formulation of quantum mechanics,[1] trajectories in space-time are assigned equal weights, and each trajectory is multiplied by the factor $\exp[(i/h)\int L\,dt]$, L is the Lagrangian of the system, and the integral is along the trajectory. The sum (integral) of the contributions of all trajectories gives the transition amplitude between the two end points.

As in the case of the spin-half system, we can define a dense sequence of measurements which check whether the particle moves along any given trajectory in space-time. This will lead to a positive answer, i.e., the particle moves along any chosen trajectory with probability one.

Consider the set of projection operators

$$\hat{\Pi}_n = |n\rangle\langle n|\,, \quad \text{for } n = 1, \ldots, N,$$

where $|n\rangle$ (in the x representation) is given by

$$|n\rangle = C\exp\{-[X - X(t_n)]^2/(2\Delta X)^2\}\exp\left[\frac{i}{\hbar}P(t_n)X\right].$$

$X(t_n)$, $P(t_n)$ are evaluated along a particular classical trajectory, at a sequence of times $t_n = n\delta t$, the total period of observation is T and $\delta t = T/N$, C is a normalizing factor, and ΔX is the uncertainty. Note that since the classical trajectory is smooth, $\dot{X}(t_n)$ and $\dot{P}(t_n)$ are well defined and so are $\delta X_n \equiv \dot{X}(t_n)\delta t$ and $\delta P_n \equiv \dot{P}(t_n)\delta t$.

We shall now prove that if the initial state is a localized wave packet around $X(t_0) \equiv X_0$, and we measure the set of operators $\hat{\Pi}_n$ (while letting $N \to \infty$), the initial state will evolve along the eigenstates of $\hat{\Pi}_n$ with probability one. We will also show that the change of the phase associated with the evolving state is well defined in terms of the sequence of operations; this change will be evaluated and shown to be equal essentially to the classical action (divided by $\hbar$).

We first note that in the infinitesimal time δt between any two measurements, the state evolves according to the Schrödinger equation and we can write the following general expression[11]:

$$|\psi(t_n + \delta t)\rangle = |n\rangle\exp\left(-\frac{i}{\hbar}\langle n|\hat{H}|n\rangle\delta t\right) + b|\psi_n^{\perp}\rangle + O(\delta t^2)\,, \tag{1}$$

where $|\psi_n^{\perp}\rangle$ is a state orthogonal to $|n\rangle$ and b is proportional to δt. If we now measure $\hat{\Pi}_{n+1}$ and get a positive answer (which is indeed the case[12]), we end up in the state $|n+1\rangle$, and the probability amplitude for this transition is

$$\langle n+1|\psi_n(t_n + \delta t)\rangle = \langle n+1|n\rangle\exp\left[-\frac{i}{\hbar}\langle n|\hat{H}|n\rangle\delta t\right] + O(\delta t)\,. \tag{2}$$

It then follows that

$$|\psi(\text{final})\rangle = |N\rangle\exp\left[-\frac{i}{\hbar}\sum_{n=1}^{N}\langle n|\hat{H}|n\rangle\delta t\right]\prod_{n=1}^{N}\langle n+1|n\rangle + O(\delta t)\,. \tag{3}$$

For a particle described by a Hamiltonian[13] $H = p^2/2m + V(X)$, the expression $\langle n|H|n\rangle$ turns out to be

$$\begin{aligned}\langle n|\hat{H}|n\rangle &= \langle n|p^2/2m + V(\hat{X})|n\rangle \\ &= \left\langle 0\left|\frac{[p - p(t_n)]^2}{2m} + V(\hat{x} - x(t_n))\right|0\right\rangle \\ &= \frac{p^2(t_n)}{2m} + V(X_n) + \theta(\Delta X^2) + \left\langle 0\left|\frac{\hat{p}^2}{2m}\right|0\right\rangle.\end{aligned} \tag{4}$$

In the limit in which $\Delta X \to 0$,[14] the above expression will simply be the classical Hamiltonian evaluated along the classical trajectory $X(t)$ plus a constant term independent of the trajectory.

The scalar product of two adjacent eigenstates in Eq. (2), gives the following contribution to the phase:

$$-\frac{i}{\hbar}\Delta p\frac{X(t_n) + X(t_{n+1})}{2} = \frac{i}{\hbar}P(t_n)\dot{X}(t_n)\delta t + \frac{i}{\hbar}[P(t_n)X(t_n) - P(t_{n+1})X(t_{n+1})]\,. \tag{5}$$

Using expressions (3)–(5), and going to the limit $N \to \infty$ ($\Sigma \to \int$), we find that the accumulated phase associated with the probability amplitude for motion along the trajectory is the sum of the familiar $(i/\hbar)\int L\,dt$, the end terms in Eq. (5), and a term independent of the trajectory ($\langle 0|p^2/2m|0\rangle T$); note that the terms proportional to $(\delta t)^2$ are negligible because $N \to \infty$ as $1/\delta t$ and therefore $N \times O(\delta t^2) \to 0$.

SETUPS THAT MEASURE THE RELATIVE PHASE OF ANY TWO TRAJECTORIES

No physical meaning can be attached to the phase assigned to an individual trajectory. However, our approach opens up the possibility of considering experimental arrangements that can measure in principle, the relative phase between any two trajectories.

We shall now write the Hamiltonian describing such a setup. This Hamiltonian has to include the degrees of freedom of the measuring device (MD) (following the well-known approach introduced by Von Neumann[15]), in order to have a full description

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

of the process as well as to prove that, indeed, the particle follows the observed trajectory with probability one:

$$\hat{H}=\hat{H}_0+\frac{(1+\hat{\sigma}_x^{(0)})}{2}\sum_n g(t-t_n)[1-\hat{\pi}_n^{(1)}]\hat{\sigma}_x^{(n)}+\frac{(1-\hat{\sigma}_x^{(0)})}{2}\sum_n g(t-t_n)[1-\hat{\pi}_n^{(2)}]\hat{\sigma}_x^{(n)}+\hat{H}_{MD}\,, \quad (6)$$

where $\hat{\pi}_n^{(1)}$ is the projection operator $|n\rangle\langle n|$, with the state $|n\rangle$ parametrized with the coordinates of the first trajectory, and $\hat{\pi}_n^{(2)}$ is similarly defined. The measuring device is described by the $\sigma_x^{(n)}$ $(n=1,2,\ldots,N)$, such that initially they are, say, in the spin-up state, and by a proper choice of the function $g(t)$ (as given afterwards), a registration (a positive result in the measurement of the projection operator $|n\rangle\langle n|$) at $t=t_n=n\delta t$ will be inferred from the fact that the nth spin did not flip.

The function $g(t)$ is defined as follows:

$$g(t)=\begin{cases} g_0 & \text{for } 0\leq t\leq\epsilon \\ 0 & \text{otherwise}\end{cases}$$

and we eventually go to the limit $\epsilon\to 0$, $g_0\to\infty$ while keeping $g_0\epsilon=\pi$. This is an impulsive measurement and therefore $\epsilon\ll\delta t$. H_0 is the Hamiltonian describing the particle and H_{MD} is the Hamiltonian of the apparatus not including the interaction term, (we obviously demand $[\sigma_i^{(n)},H_{MD}]\neq 0$ and $[\sigma_i^{(0)},H_{MD}]=0$, $i=x,y,z$).

$\hat{\sigma}_x^{(0)}$ is the operator describing the spin of a particle in the measuring device that, according to its initial value, the process of dense measurements will evolve, i.e., if $\sigma_x^{(0)}=+1$ the process will evolve along trajectory one, if $\sigma_x^{(0)}=-1$, along trajectory two, and if in a superposition of the two state, the process of measurements will also be a superposition of two processes, in the sense described below. We will show that if the initial state of the whole system is given by

$$|\psi(0)\rangle=\frac{c}{\sqrt{2}}\exp[-(x-x_0)^2/(2\Delta x)^2]\left[\exp\left(\frac{i}{\hbar}xp_0\right)|\sigma_x^{(0)}=+1\rangle\exp\left(-\frac{i}{\hbar}x_0p_0\right)+\exp\left(\frac{i}{\hbar}xp_0'\right)|\sigma_x^{(0)}=-1\rangle\exp\left(-\frac{i}{\hbar}x_0p_0'\right)\right]\prod_{n=1}^{N}|\sigma_x^{(n)}=+1\rangle$$

$$\simeq c\exp[-(x-x_0)^2/(2\Delta x)^2]|\sigma_x^{(0)}=+1\rangle\prod_{n=1}^{N}|\sigma_x^{(n)}=+1\rangle \quad (7a)$$

[the last (approximate) equality is due to $\Delta x\to 0$; remember that we always go to this limit[14]], then the final state will be

$$|\psi(\text{final})\rangle\simeq\frac{c}{\sqrt{2}}\exp\{-[x-x(t_n)]^2/(2\Delta x)^2\}\left[|\sigma_x^{(0)}=+1\rangle+\exp\left(\frac{i}{\hbar}\alpha\right)|\sigma_x^{(0)}=-1\rangle\right]\prod_{n=1}^{N}|\sigma_x^{(n)}=+1\rangle, \quad (7b)$$

where $\alpha\equiv\int_1 L\,dt-\int_2 L\,dt$; this means that the direction of the zeroth spin is sensitive to α.

Let us outline the steps that lead to the above result: First, assume that the initial state (of the whole system) is the one corresponding to the measurement of the first trajectory; then we find that the state evolves as follows:

(a) In the infinitesimal time δt before the first measurement we end up with [see Eq. (1)]

$$|\psi(t)\rangle=\left[|0\rangle\exp\left(-\frac{i}{\hbar}\langle 0|H|0\rangle\delta t-\frac{i}{\hbar}x_0P_0\right)+b|\psi_0^\perp\rangle+O(\delta t^2)\right]|\text{MD}^{(0)}\rangle,$$

where $|\text{MD}^{(0)}\rangle$ means no spins have been flipped in the measuring device, etc.

(b) At $t=\delta t$ we have an impulsive measurement (i.e., only the interaction with the MD is relevant[15]). This changes $|\psi(t)\rangle$ into

$$|\psi(t=\delta t)\rangle=\left[\langle 1|0\rangle\exp\left(-\frac{i}{\hbar}\langle 0|H|0\rangle\delta t-\frac{i}{\hbar}x_0P_0\right)+O(\delta t^2)\right]|1\rangle|\text{MD}^{(0)}\rangle+\left[\langle\psi_1^\perp|0\rangle\exp\left(-\frac{i}{\hbar}\langle 0|H|0\rangle\delta t-x_0P_0\right)+O(\delta t)\right]|\psi_1^\perp\rangle|\text{MD}^{(1)}\rangle,$$

where $|\psi_1^\perp\rangle$ is a state orthogonal to $|1\rangle$.

This procedure can be repeated, i.e., the free Hamiltonian acts on the state (in between measurement and then the interaction term (which is impulsive) takes over, etc., N times.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Letting $N\to\infty$ and $\delta t\to 0$ while keeping $N\delta t = T$ we realize that the *probability* that one of the spins had been flipped (in the measuring device) is a sum of N terms each proportional to $(\delta t)^2$ (and higher orders), so that $N(\delta t)^2 \to_{N\to\infty} 0$. This is so because of the fact that terms in the amplitude that are proportional to δt are multiplied by orthogonal states of the measuring device. Then, using Eqs. (3)–(5) we can write the final state as follows:

$$|\psi(\text{final})\rangle = |N\rangle\,|\sigma_x^{(0)}=+1\rangle \exp\left(\frac{i}{\hbar}\int_1 L\,dt\right)\exp\left(-\frac{i}{\hbar}x_N P_N\right)\exp\left(-\frac{i}{\hbar}\langle 0|\hat{P}^2/2m|0\rangle T\right)\prod_{n=1}^{N}|\sigma_x^{(n)}=+1\rangle .$$

If we superpose two such states coming from different trajectories and recall that $\Delta X\to 0$ (which will make the end term disappear), we end up with the final state given by Eq. (7b), which means we also observe the relative phase between any two trajectories.

CONCLUSIONS

In this paper we have shown that it is possible to define trajectories in Hilbert space connecting any two states. If one measures successively a dense set of projection operators corresponding to the states that define a trajectory,[16] one moves with probability one from the initial state to the final state through the whole set.

We pointed out that the problem of freezing a state as a result of continuous observations is a particular case of this more general notion of trajectory. We have applied the above notion to give an observable meaning to an individual Feynman path, and we have shown that a phase is systematically accumulated during this process of observation.

An experimental arrangement was suggested in order to measure the relative phase between any two trajectories; this phase was shown to agree with the one postulated by Feynman. The concepts of trajectories and their associated phases can serve as the starting point for an extension of the Feynman formalism to include more general Hamiltonians (see the Appendix). These ideas will be further discussed in a future publication.

We thank Professor E. Lerner for helpful comments.

APPENDIX

Consider, for example, the following trajectory that does not have a classical counterpart. We use the eigenoperators of the two-slit experiment[8] which are defined as follows:

$$\bar{\sigma}_3 = \frac{\sin(\pi\hat{x}/l)}{|\sin(\pi\hat{x}/l)|} ,$$

$$\bar{\sigma}_1 = \cos(\hat{p}l/h) - \sin(\hat{p}l/h)\bar{\sigma}_3 ,$$

$$\bar{\sigma}_2 = \sin(\hat{p}l/h) + i\cos(\hat{p}l/h)\bar{\sigma}_3 ,$$

where $\hat{x}$ is the position operator along the line connecting any two given points in space, and the origin $(x=0)$ is in the middle of the two.

Assuming that the initial state $|0\rangle$ and final state $|N\rangle$ of the previous example are localized wave packets with negligible overlap, we can define a trajectory connecting them, as follows:

Let us call $|0\rangle$ and $|N\rangle$ the "spin-up" and "spin-down" states of $\bar{\sigma}_3$, respectively. Then, if we perform the sequence of measurements $\bar{\sigma}_{\alpha_n} = \bar{\sigma}_3\cos\alpha_n + \sigma_2\sin\alpha_n$, $n = 0, 1, \ldots, N$, where $\alpha_n = n\pi/N$ and let $N\to\infty$, we find that the initial state $|0\rangle$ will evolve along the trajectory defined by the eigenoperators of $\bar{\sigma}_{\alpha_n}$, finally reaching the state $|N\rangle$.[17] The proof goes in complete analogy to the one we gave for the spin-half system. Note that in this last example the Hamiltonian includes a nonlocal term.[18]

[1] R. P. Feynman, Rev. Mod. Phys. 20, 367 (1948).
[2] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, New York, 1974), p. 7.
[3] Ch. N. Friedman, Ann. Phys. (N.Y.) 98, 87 (1976).
[4] B. Mirsa and E. C. G. Sudarshan, J. Math. Phys. 18, 756 (1977).
[5] I. Bloch and D. A. Burba, Phys. Rev. D 10, 3206 (1974).
[6] G. R. Allcock, Ann. Phys. (N.Y.) 53, 253 (1969).
[7] (a) E. Eberle, Lett. Nuovo Cimento 20, 272 (1971); (b) H. Elstein and A. J. F. Siegerert, Ann. Phys. (N.Y.) 68, 508 (1971). More relevant references can be found in the above-mentioned articles (Refs. 3–6).
[8] Y. Aharonov *et al.*, Int. J. Theor. Phys. 3, 443 (1970).
[9] Note that in the case of the decaying system, the deterministic measurements have to involve also the field degrees of freedom.
[10] This is equivalent to the unitary transformation $U = \exp[(i/\hbar)(\omega t\hat{\sigma}_x/2)]$ where $H = \vec{\mu}\cdot\vec{B} = \hbar\omega\hat{\sigma}_x/2$.
[11] This is evident from the following relations:

$$\psi(t) = \psi(0) + \dot{\psi}(0)\delta t + O(\delta t^2),$$

$\dot{\psi}(0) = (1/i\hbar)\hat{H}\psi(0)$ (Schrödinger equation at $t=0$); by in-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

serting (b) into (a) we get

$$\psi(t) = \psi(0) + (1/i\hbar)\hat{H}\psi(0)\delta t + O(\delta t^2),$$

$$\langle\psi(0)|\psi(t)\rangle = 1 + (1/i\hbar)\langle 0|\hat{H}|0\rangle\delta t + O(\delta t^2)$$

$$= \exp[-(i/\hbar)\langle 0|\hat{H}|0\rangle\delta t] + O(\delta t^2).$$

Now we write an alternative identity for $|\psi(t)\rangle$, namely,

$$\psi(t) = a\psi(0) + b|\psi_0^\perp\rangle,$$

where $a \equiv \exp[(i/\hbar)\langle 0|H|0\rangle\delta t] + O(\delta t^2)$ and $|\psi_0^\perp\rangle$ is an orthogonal state to $\psi(0)$. Finally we note that $b \sim \delta t$ since

$$1 \equiv \langle\psi(t)|\psi(t)\rangle = |a|^2\langle\psi_0|\psi_0\rangle + \langle\psi_0^\perp|\psi_0^\perp\rangle|b|^2$$

and because $\langle\psi(0)|\psi(0)\rangle \equiv 1$ and $\langle\psi_0^\perp|\psi_0^\perp\rangle = 1$, $|b|^2 = 1 - |a|^2 = O(\delta t^2)$.

[12]See a detailed analysis of the measurement in the last section of the paper, but note that a straightforward proof can be given in analogy to the spin case, where the only differences are the following: (a) Instead of the magnetic field we now have a more general Hamiltonian, but still the probability that the state escapes from the original one is of the order $(\delta t)^2$ (more exactly, $|\langle n|t\rangle|^2 = 1 - [(\delta t)^2/\hbar^2](\Delta E)^2 + O(\delta t^3)$, where ΔE is the uncertainty of the energy in the state $|n\rangle$ and $|t\rangle$ is the state between t_n and t_{n+1}). (b) Instead of measuring the same projection operator at intervals δt, we now measure projection operators of states that are "slightly" shifted, namely, $|\langle n+1|n\rangle|^2 \sim (\delta t)^2$. It is easy to show now that $|\langle n+1|t\rangle|^2 \sim (\delta t)^2$, and from this point the argument follows the one given in the spin case.

[13]It can be generalized straightforwardly to the three dimensional case including a vector potential.

[14]Recall that we demanded $\prod_{n=1}^N |\langle n+1|n\rangle^2 \to 1$ as $N\to\infty$, which is satisfied if $[(\delta x)_{\max}/\Delta x]^2 N \to 0$ as $N\to\infty$ (because $|\langle n+1|n\rangle|^2 \sim \exp[-(\delta x/\Delta x)^2]$ and $(\delta x)_{\max}$ is the maximum distance between the centers of two adjacent Gaussians along the trajectory), then using the fact that $N \equiv T/\delta t$ and $(\delta x)_{\max} = \dot{x}_{\max}\,\delta t$ we get

$$\frac{T(\dot{X})_{\max}^2\delta t}{(\Delta x)^2} \xrightarrow[\delta t\to 0]{} 0.$$

This means that we first have to go to the limit $\delta t \to 0$ and then to choose Δx arbitrarily small.

[15]See, for example, Y. Aharonov and J. L. Safko, Ann. Phys. (N. Y.) 91, 279 (1975).

[16]In the sense that if a total of N measurements are made (in some given period of time T) then the probability of success in each measurement has to be at least of the order $(1 - 1/N^2)$, and $N \gg 1$.

[17]The particle makes quantum jumps from its initial position to the final position without going through intermediate ones.

[18]The realization of these types of Hamiltonians is considered in a recent paper by Y. Aharonov and E. Lerner, Phys. Rev. D 20, 1877 (1979).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

THE

PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 115, NO. 3 — AUGUST 1, 1959

Significance of Electromagnetic Potentials in the Quantum Theory

Y. AHARONOV AND D. BOHM
H. H. Wills Physics Laboratory, University of Bristol, Bristol, England
(Received May 28, 1959; revised manuscript received June 16, 1959)

In this paper, we discuss some interesting properties of the electromagnetic potentials in the quantum domain. We shall show that, contrary to the conclusions of classical mechanics, there exist effects of potentials on charged particles, even in the region where all the fields (and therefore the forces on the particles) vanish. We shall then discuss possible experiments to test these conclusions; and, finally, we shall suggest further possible developments in the interpretation of the potentials.

1. INTRODUCTION

In classical electrodynamics, the vector and scalar potentials were first introduced as a convenient mathematical aid for calculating the fields. It is true that in order to obtain a classical canonical formalism, the potentials are needed. Nevertheless, the fundamental equations of motion can always be expressed directly in terms of the fields alone.

In the quantum mechanics, however, the canonical formalism is necessary, and as a result, the potentials cannot be eliminated from the basic equations. Nevertheless, these equations, as well as the physical quantities, are all gauge invariant; so that it may seem that even in quantum mechanics, the potentials themselves have no independent significance.

In this paper, we shall show that the above conclusions are not correct and that a further interpretation of the potentials is needed in the quantum mechanics.

2. POSSIBLE EXPERIMENTS DEMONSTRATING THE ROLE OF POTENTIALS IN THE QUANTUM THEORY

In this section, we shall discuss several possible experiments which demonstrate the significance of potentials in the quantum theory. We shall begin with a simple example.

Suppose we have a charged particle inside a "Faraday cage" connected to an external generator which causes the potential on the cage to alternate in time. This will add to the Hamiltonian of the particle a term $V(x,t)$ which is, for the region inside the cage, a function of time only. In the nonrelativistic limit (and we shall assume this almost everywhere in the following discussions) we have, for the region inside the cage, $H=H_0+V(t)$ where H_0 is the Hamiltonian when the generator is not functioning, and $V(t)=e\phi(t)$. If $\psi_0(x,t)$ is a solution of the Hamiltonian H_0, then the solution for H will be

$$\psi=\psi_0 e^{-iS/\hbar},\quad S=\int V(t)dt,$$

which follows from

$$i\hbar\frac{\partial\psi}{\partial t}=\left(i\hbar\frac{\partial\psi_0}{\partial t}+\psi_0\frac{\partial S}{\partial t}\right)e^{-iS/\hbar}=[H_0+V(t)]\psi=H\psi.$$

The new solution differs from the old one just by a phase factor and this corresponds, of course, to no change in any physical result.

Now consider a more complex experiment in which a single coherent electron beam is split into two parts and each part is then allowed to enter a long cylindrical metal tube, as shown in Fig. 1.

After the beams pass through the tubes, they are combined to interfere coherently at F. By means of time-determining electrical "shutters" the beam is chopped into wave packets that are long compared with the wavelength λ, but short compared with the length of the tubes. The potential in each tube is determined by a time delay mechanism in such a way that the potential is zero in region I (until each packet is well inside its tube). The potential then grows as a function of time, but differently in each tube. Finally, it falls back to zero, before the electron comes near the

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

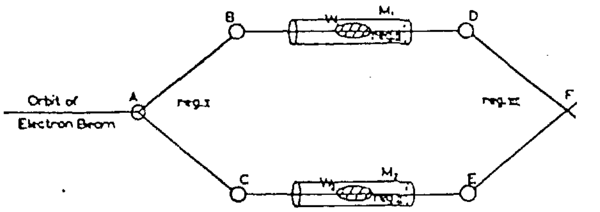


FIG. 1. Schematic experiment to demonstrate interference with time-dependent scalar potential. A, B, C, D, E: suitable devices to separate and divert beams. W_1, W_2: wave packets. M_1, M_2: cylindrical metal tubes. F: interference region.

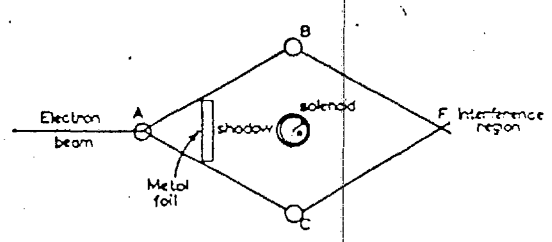


FIG. 2. Schematic experiment to demonstrate interference with time-independent vector potential.

other edge of the tube. Thus the potential is nonzero only while the electrons are well inside the tube (region II). When the electron is in region III, there is again no potential. The purpose of this arrangement is to ensure that the electron is in a time-varying potential without ever being in a field (because the field does not penetrate far from the edges of the tubes, and is nonzero only at times when the electron is far from these edges).

Now let $\psi(x,t)=\psi_1^0(x,t)+\psi_2^0(x,t)$ be the wave function when the potential is absent (ψ_1^0 and ψ_2^0 representing the parts that pass through tubes 1 and 2, respectively). But since V is a function only of t wherever ψ is appreciable, the problem for each tube is essentially the same as that of the Faraday cage. The solution is then

$$\psi=\psi_1^0e^{-iS_1/\hbar}+\psi_2^0e^{-iS_2/\hbar},$$

where

$$S_1=e\int\varphi_1dt,\quad S_2=e\int\varphi_2dt.$$

It is evident that the interference of the two parts at F will depend on the phase difference $(S_1-S_2)/\hbar$. Thus, there is a physical effect of the potentials even though no force is ever actually exerted on the electron. The effect is evidently essentially quantum-mechanical in nature because it comes in the phenomenon of interference. We are therefore not surprised that it does not appear in classical mechanics.

From relativistic considerations, it is easily seen that the covariance of the above conclusion demands that there should be similar results involving the vector potential, A.

The phase difference, $(S_1-S_2)/\hbar$, can also be expressed as the integral $(e/\hbar)\oint\varphi dt$ around a closed circuit in space-time, where φ is evaluated at the place of the center of the wave packet. The relativistic generalization of the above integral is

$$\frac{e}{\hbar}\oint\left(\varphi dt-\frac{\mathbf{A}}{c}\cdot d\mathbf{x}\right),$$

where the path of integration now goes over any closed circuit in space-time.

As another special case, let us now consider a path in space only (t=constant). The above argument suggests that the associated phase shift of the electron wave function ought to be

$$\Delta S/\hbar=-\frac{e}{c\hbar}\oint\mathbf{A}\cdot d\mathbf{x},$$

where $\oint\mathbf{A}\cdot d\mathbf{x}=\int\mathbf{H}\cdot d\mathbf{s}=\phi$ (the total magnetic flux inside the circuit).

This corresponds to another experimental situation. By means of a current flowing through a very closely wound cylindrical solenoid of radius R, center at the origin and axis in the z direction, we create a magnetic field, H, which is essentially confined within the solenoid. However, the vector potential, A, evidently, cannot be zero everywhere outside the solenoid, because the total flux through every circuit containing the origin is equal to a constant

$$\phi_0=\int\mathbf{H}\cdot d\mathbf{s}=\int\mathbf{A}\cdot d\mathbf{x}.$$

To demonstrate the effects of the total flux, we begin, as before, with a coherent beam of electrons. (But now there is no need to make wave packets.) The beam is split into two parts, each going on opposite sides of the solenoid, but avoiding it. (The solenoid can be shielded from the electron beam by a thin plate which casts a shadow.) As in the former example, the beams are brought together at F (Fig. 2).

The Hamiltonian for this case is

$$H=\frac{[\mathbf{P}-(e/c)\mathbf{A}]^2}{2m}.$$

In singly connected regions, where $\mathbf{H}=\nabla\times\mathbf{A}=0$, we can always obtain a solution for the above Hamiltonian by taking $\psi=\psi_0e^{-iS/\hbar}$, where ψ_0 is the solution when $\mathbf{A}=0$ and where $\nabla S/\hbar=(e/c)\mathbf{A}$. But, in the experiment discussed above, in which we have a multiply connected region (the region outside the solenoid), $\psi_0e^{-iS/\hbar}$ is a non-single-valued function[1] and therefore, in general, not a permissible solution of Schrödinger's equation. Nevertheless, in our problem it is still possible to use such solutions because the wave function splits into two parts $\psi=\psi_1+\psi_2$, where ψ_1 represents the beam on

[1] Unless $\phi_0=nhc/e$, where n is an integer.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

one side of the solenoid and ψ_2 the beam on the opposite side. Each of these beams stays in a simply connected region. We therefore can write

$$\psi_1=\psi_1^0 e^{-iS_1/\hbar}, \quad \psi_2=\psi_2^0 e^{-iS_2/\hbar},$$

where S_1 and S_2 are equal to $(e/c)\int \mathbf{A}\cdot d\mathbf{x}$ along the paths of the first and second beams, respectively. (In Sec. 4, an exact solution for this Hamiltonian will be given, and it will confirm the above results.)

The interference between the two beams will evidently depend on the phase difference,

$$(S_1-S_2)/\hbar=(e/\hbar c)\oint \mathbf{A}\cdot d\mathbf{x}=(e/\hbar c)\phi_0.$$

This effect will exist, even though there are no magnetic forces acting in the places where the electron beam passes.

In order to avoid fully any possible question of contact of the electron with the magnetic field we note that our result would not be changed if we surrounded the solenoid by a potential barrier that reflects the electrons perfectly. (This, too, is confirmed in Sec. 4.)

It is easy to devise hypothetical experiments in which the vector potential may influence not only the interference pattern but also the momentum. To see this, consider a periodic array of solenoids, each of which is shielded from direct contact with the beam by a small plate. This will be essentially a grating. Consider first the diffraction pattern without the magnetic field, which will have a discrete set of directions of strong constructive interference. The effect of the vector potential will be to produce a shift of the relative phase of the wave function in different elements of the gratings. A corresponding shift will take place in the directions, and therefore the momentum of the diffracted beam.

3. A PRACTICABLE EXPERIMENT TO TEST FOR THE EFFECTS OF A POTENTIAL WHERE THERE ARE NO FIELDS

As yet no direct experiments have been carried out which confirm the effect of potentials where there is no field. It would be interesting therefore to test whether such effects actually exist. Such a test is, in fact, within the range of present possibilities.[2] Recent experiments[3,4] have succeeded in obtaining interference from electron beams that have been separated in one case by as much as 0.8 mm.[3] It is quite possible to wind solenoids which are smaller than this, and therefore to place them between the separate beams. Alternatively, we may obtain localized lines of flux of the right magnitude (the magnitude has to be of the order of $\phi_0=2\pi c\hbar/e\sim 4\times 10^{-7}$ gauss cm^2) by means of fine permanently magnetized "whiskers".[5] The solenoid can be used in Marton's device,[3] while the whisker is suitable for another experimental setup[4] where the separation is of the order of microns and the whiskers are even smaller than this.

In principle, we could do the experiment by observing the interference pattern with and without the magnetic flux. But since the main effect of the flux is only to displace the line pattern without changing the interval structure, this would not be a convenient experiment to do. Instead, it would be easier to vary the magnetic flux within the same exposure for the detection of the interference patterns. Such a variation would, according to our previous discussion, alter the sharpness and the general form of the interference bands. This alteration would then constitute a verification of the predicted phenomena.

When the magnetic flux is altered, there will, of course, be an induced electric field outside the solenoid, but the effects of this field can be made negligible. For example, suppose the magnetic flux were suddenly altered in the middle of an exposure. The electric field would then exist only for a very short time, so that only a small part of the beam would be affected by it.

4. EXACT SOLUTION FOR SCATTERING PROBLEMS

We shall now obtain an exact solution for the problem of the scattering of an electron beam by a magnetic field in the limit where the magnetic field region tends to a zero radius, while the total flux remains fixed. This corresponds to the setup described in Sec. 2 and shown in Fig. 2. Only this time we do not split the plane wave into two parts. The wave equation outside the magnetic field region is, in cylindrical coordinates,

$$\left[\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{1}{r^2}\left(\frac{\partial}{\partial\theta}+i\alpha\right)^2+k^2\right]\psi=0, \qquad (1)$$

where $\mathbf{k}$ is the wave vector of the incident particle and $\alpha=-e\phi/ch$. We have again chosen the gauge in which $A_r=0$ and $A_\theta=\phi/2\pi r$.

The general solution of the above equation is

$$\psi=\sum_{m=-\infty}^{\infty} e^{im\theta}[a_m J_{m+\alpha}(kr)+b_m J_{-(m+\alpha)}(kr)], \qquad (2)$$

where a_m and b_m are arbitrary constants and $J_{m+\alpha}(kr)$ is a Bessel function, in general of fractional order (dependent on ϕ). The above solution holds only for $r>$ [illegible] $r<R$ (inside the magnetic field) the solution has been worked out.[6] By matching the solutions at $r=R$ it is easily shown that only Bessel functions of positive order will remain, when R approaches zero.

[2] Dr. Chambers is now making a preliminary experimental study of this question at Bristol.

[3] L. Marton, Phys. Rev. 85, 1057 (1952); 90, 490 (1953). Marton, Simpson, and Suddeth, Rev. Sci. Instr. 25, 1099 (1954).

[4] G. Mollenstedt, Naturwissenschaften 42, 41 (1955); G. Mollenstedt and H. Düker, Z. Physik 145, 377 (1956).

[5] See, for example, Sidney S. Brenner, Acta Met. 4, 62 (1956).

[6] L. Page, Phys. Rev. 36, 444 (1930).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

This means that the probability of finding the particle inside the magnetic field region approaches zero with R. It follows that the wave function would not be changed if the electron were kept away from the field by a barrier whose radius also went to zero with R.

The general solution in the limit of R tending to zero is therefore

$$\psi=\sum_{m=-\infty}^{\infty} a_m J_{|m+\alpha|} e^{im\theta}. \quad (3)$$

We must then choose a_m so that ψ represents a beam of electrons that is incident from the right ($\theta=0$). It is important, however, to satisfy the initial condition that the current density,

$$\mathbf{j}=\frac{\hbar(\psi^*\nabla\psi-\psi\nabla\psi^*)}{2im}-\frac{e}{mc}\mathbf{A}\psi^*\psi, \quad (4)$$

shall be constant and in the x direction. In the gauge that we are using, we easily see that the correct incident wave is $\psi_{\text{inc}}=e^{-ikx}e^{-i\alpha\theta}$. Of course, this wave function holds only to the right of the origin, so that no problem of multiple-valuedness arises.

We shall show in the course of this calculation that the above conditions will be satisfied by choosing $a_m=(-i)^{|m+\alpha|}$, in which case, we shall have

$$\psi=\sum_{m=-\infty}^{\infty}(-i)^{|m+\alpha|}J_{|m+\alpha|}e^{im\theta}.$$

It is convenient to split ψ into the following three parts: $\psi=\psi_1+\psi_2+\psi_3$, where

$$\psi_1=\sum_{m=1}^{\infty}(-i)^{m+\alpha}J_{m+\alpha}e^{im\theta},$$

$$\psi_2=\sum_{m=-\infty}^{-1}(-i)^{m+\alpha}J_{m+\alpha}e^{im\theta},$$

$$=\sum_{m=1}^{\infty}(-i)^{m-\alpha}J_{m-\alpha}e^{-im\theta}, \quad (5)$$

$$\psi_3=(-i)^{|\alpha|}J_{|\alpha|}.$$

Now ψ_1 satisfies the simple differential equation

$$\frac{\partial\psi_1}{\partial r'}=\sum_{m=1}^{\infty}(-i)^{m+\alpha}J_{m+\alpha}'e^{im\theta}$$

$$=\sum_{m=1}^{\infty}(-i)^{m+\alpha}\frac{J_{m+\alpha-1}-J_{m+\alpha+1}}{2}e^{im\theta}, \quad r'=kr \quad (6)$$

where we have used the well-known formula for Bessel functions:

$$dJ_\gamma(r)/dr=\tfrac{1}{2}(J_{\gamma-1}-J_{\gamma+1}).$$

As a result, we obtain

$$\frac{\partial\psi_1}{\partial r'}=\frac{1}{2}\sum_{m'=0}^{\infty}(-i)^{m'+\alpha+1}J_{m'+\alpha}e^{i(m'+1)\theta}-\frac{1}{2}\sum_{m'=2}^{\infty}(-i)^{m'+\alpha-1}J_{m'+\alpha}e^{i(m'-1)\theta}$$

$$=\frac{1}{2}\sum_{m'=1}^{\infty}(-i)^{m'+\alpha}J_{m'+\alpha}e^{im'\theta}(-ie^{i\theta}+i^{-1}e^{-i\theta})+\tfrac{1}{2}(-i)^{\alpha}[J_{\alpha+1}-ie^{i\theta}J_{\alpha}]. \quad (7)$$

So

$$\partial\psi_1/\partial r'=-i\cos\theta\psi_1+\tfrac{1}{2}(-i)^{\alpha}(J_{\alpha+1}-iJ_\alpha e^{i\theta}).$$

This differential equation can be easily integrated to give

$$\psi_1=A\int_0^{r'}e^{ir'\cos\theta}[J_{\alpha+1}-iJ_\alpha e^{i\theta}]dr', \quad (8)$$

where

$$A=\tfrac{1}{2}(-i)^{\alpha}e^{-ir'\cos\theta}.$$

The lower limit of the integration is determined by the requirement that when r' goes to zero, ψ_1 also goes to to zero because, as we have seen, ψ_1 includes Bessel functions of positive order only.

In order to discuss the asymptotic behavior of ψ_1, let us write it as $\psi_1=A[I_1-I_2]$, where

$$I_1=\int_0^{\infty}e^{ir'\cos\theta}[J_{\alpha+1}-ie^{i\theta}J_\alpha]dr',$$

$$I_2=\int_r^{\infty}e^{ir'\cos\theta}[J_{\alpha+1}-ie^{i\theta}J_\alpha]dr'. \quad (9)$$

The first of these integrals is known[7]:

$$\int_0^{\infty}e^{i\beta r}J_\alpha(kr)=\frac{e^{i[\alpha\arcsin(\beta/k)]}}{(k^2-\beta^2)^{\frac{1}{2}}}, \quad 0<\beta<k, \quad -2<\alpha.$$

In our cases, $\beta=\cos\theta$, $k=1$, so that

$$I_1=\left[\frac{e^{i\alpha(\frac{1}{2}\pi-|\theta|)}}{|\sin\theta|}-ie^{i\theta}\frac{e^{i(\alpha+1)(\frac{1}{2}\pi-|\theta|)}}{|\sin\theta|}\right]. \quad (10)$$

Because the integrand is even in θ, we have written the final expression for the above integral as a function of $|\theta|$ and of $|\sin\theta|$. Hence

$$I_1=e^{i\alpha(\frac{1}{2}\pi-|\theta|)}\left[\frac{ie^{-i|\theta|}-ie^{i\theta}}{|\sin\theta|}\right]$$

$$=0 \quad \text{for } \theta<0,$$

$$=e^{-i\alpha\theta}2i^{\alpha} \quad \text{for } \theta>0, \quad (11)$$

where we have taken θ as going from $-\pi$ to π.

[7] See, for example, W. Gröbner and N. Hofreiter, *Integraltafel* (Springer-Verlag, Berlin, 1949).

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

We shall see presently that I_1 represents the largest term in the asymptotic expansion of ψ_1. The fact that it is zero for $\theta<0$ shows that this part of ψ_1 passes (asymptotically) only on the upper side of the singularity. To explain this, we note that ψ_1 contains only positive values of m, and therefore of the angular momentum. It is quite natural then that this part of ψ_1 goes on the upper side of the singularity. Similarly, since according to (5)

$$\psi_2(r',\theta,\alpha)=\psi_1(r',-\theta,-\alpha),$$

it follows that ψ_2 will behave oppositely to ψ_1 in this regard, so that together they will make up the correct incident wave.

Now, in the limit of $r'\rightarrow\infty$ we are allowed to take in the integrand of I_2 the first asymptotic term of J_α,[8] namely $J_\alpha\rightarrow(2/\pi r)^{\frac{1}{2}}\cos(r'-\frac{1}{2}\alpha-\frac{1}{4}\pi)$. We obtain

$$I_2=\int_r^\infty e^{ir'\cos\theta}(J_{\alpha+1}-ie^{i\theta}J_\alpha)dr'\rightarrow C+D, \quad (12)$$

where

$$C=\int_r^\infty e^{ir'\cos\theta}[\cos(r'-\tfrac{1}{2}(\alpha+1)\pi-\tfrac{1}{4}\pi)]\frac{dr'}{(r')^{\frac{1}{2}}}\left(\frac{2}{\pi}\right)^{\frac{1}{2}},$$
$$D=\int_r^\infty e^{ir'\cos\theta}[\cos(r'-\tfrac{1}{2}\alpha-\tfrac{1}{4}\pi)]\frac{dr'}{(r')^{\frac{1}{2}}}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}(-i)e^{i\theta}. \quad (13)$$

Then

$$C=\int_r^\infty e^{ir'\cos\theta}\Big[e^{i[r'-\frac{1}{2}(\alpha+1)\pi-\frac{1}{4}\pi]}+e^{-i[r'-\frac{1}{2}(\alpha+1)\pi-\frac{1}{4}\pi]}\Big]\frac{dr'}{(2\pi r')^{\frac{1}{2}}}$$
$$=\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\frac{(-i)^{\alpha+\frac{1}{2}}}{(1+\cos\theta)^{\frac{1}{2}}}\int_{[r'(1+\cos\theta)]^{\frac{1}{2}}}^\infty\exp(+iz^2)dz$$
$$+\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\frac{i^{\alpha+\frac{1}{2}}}{(1-\cos\theta)^{\frac{1}{2}}}\int_{[r'(1-\cos\theta)]^{\frac{1}{2}}}^\infty\exp(-iz^2)dz, \quad (14)$$

where we have put

$$z=[r'(1+\cos\theta)]^{\frac{1}{2}} \quad\text{and}\quad z=[r'(1-\cos\theta)]^{\frac{1}{2}},$$

respectively.

Using now the well-known asymptotic behavior of the error function,[9]

$$\int_a^\infty\exp(iz^2)dz\rightarrow\frac{i}{2}\frac{\exp(ia^2)}{a},$$
$$\int_a^\infty\exp(-iz^2)dz\rightarrow\frac{-i}{2}\frac{\exp(-ia^2)}{a}, \quad (15)$$

we finally obtain

$$C=\left[\frac{(-i)^{\alpha+\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\frac{e^{ir'}}{[r'(1+\cos\theta)^2]^{\frac{1}{2}}}+\frac{i^{\alpha+\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\frac{e^{-ir'}}{[r'(1-\cos\theta)^2]^{\frac{1}{2}}}\right]e^{ir'\cos\theta}, \quad (16)$$

$$D=\left[\frac{(-i)^{\alpha-\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\frac{e^{ir'}}{[r'(1+\cos\theta)^2]^{\frac{1}{2}}}+\frac{i^{\alpha-\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\frac{e^{-ir'}}{[r'(1-\cos\theta)^2]^{\frac{1}{2}}}\right]e^{ir'\cos\theta}(-i)e^{i\theta}. \quad (17)$$

Now adding (16) and (17) together and using (13) and (9), we find that the term of $1/(r')^{\frac{1}{2}}$ in the asymptotic expansion of ψ_1 is

$$\frac{(-i)^{\frac{1}{2}}}{2(2\pi)^{\frac{1}{2}}}\left[(-1)^\alpha\frac{e^{ir'}}{(r')^{\frac{1}{2}}}\frac{1+e^{i\theta}}{1+\cos\theta}+i\frac{e^{-ir'}}{(r')^{\frac{1}{2}}}\frac{1-e^{i\theta}}{1-\cos\theta}\right]. \quad (18)$$

Using again the relation between ψ_1 and ψ_2 we obtain for the corresponding term in ψ_2

$$\frac{(-i)^{\frac{1}{2}}}{2(2\pi)^{\frac{1}{2}}}\left[(-1)^{-\alpha}\frac{e^{ir'}}{(r')^{\frac{1}{2}}}\frac{1+e^{-i\theta}}{1+\cos\theta}+i\frac{e^{-ir'}}{(r')^{\frac{1}{2}}}\frac{1-e^{-i\theta}}{1-\cos\theta}\right]. \quad (19)$$

Adding (18) and (19) and using (11), we finally get

$$\psi_1+\psi_2\rightarrow\frac{(-i)^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}}\left[\frac{ie^{-ir'}}{(r')^{\frac{1}{2}}}+\frac{e^{ir'}}{(r')^{\frac{1}{2}}}\frac{\cos(\pi\alpha-\frac{1}{2}\theta)}{\cos(\frac{1}{2}\theta)}\right]+e^{-i(r'\cos\theta+\alpha\theta)}. \quad (20)$$

There remains the contribution of ψ_3, whose asymptotic behavior is [see Eq. (12)]

$$(-i)^{|\alpha|}J_{|\alpha|}(r')\rightarrow(-i)^{|\alpha|}\left(\frac{2}{\pi r'}\right)^{\frac{1}{2}}\cos(r'-\tfrac{1}{4}\pi-\tfrac{1}{2}|\alpha|\pi).$$

Collecting all terms, we find

$$\psi=\psi_1+\psi_2+\psi_3\rightarrow e^{-i(\alpha\theta+r'\cos\theta)}+\frac{e^{ir'}}{(2\pi ir')^{\frac{1}{2}}}\sin\pi\alpha\frac{e^{\pm i\theta/2}}{\cos(\theta/2)}, \quad (21)$$

where the $\pm$ sign is chosen according to the sign of α.

The first term in equation (21) represents the incident wave, and the second the scattered wave.[10] The scattering cross section is therefore

$$\sigma=\frac{\sin^2\pi\alpha}{2\pi}\frac{1}{\cos^2(\theta/2)}. \quad (22)$$

[8] E. Jahnke and F. Emde, *Tables of Functions* (Dover Publications, Inc., New York, 1943), fourth edition, p. 138.
[9] Reference 8, p. 24.
[10] In this way, we verify, of course, that our choice of the a_n for Eq. (3) satisfies the correct boundary conditions.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

When $\alpha=n$, where n is an integer, then σ vanishes. This is analogous to the Ramsauer effect.[11] σ has a maximum when $\alpha=n+\frac{1}{2}$.

The asymptotic formula (21) holds only when we are not on the line $\theta=\pi$. The exact solution, which is needed on this line, would show that the second term will combine with the first to make a single-valued wave function, despite the non-single-valued character of the two parts, in the neighborhood of $\theta=\pi$. We shall see this in more detail presently for the special case $\alpha=n+\frac{1}{2}$.

In the interference experiment discussed in Sec. 2, diffraction effects, represented in Eq. (21) by the scattered wave, have been neglected. Therefore, in this problem, it is adequate to use the first term of Eq. (21). Here, we see that the phase of the wave function has a different value depending on whether we approach the line $\theta=\pm\pi$ from positive or negative angles, i.e., from the upper or lower side. This confirms the conclusions obtained in the approximate treatment of Sec. 2.

We shall discuss now the two special cases that can be solved exactly. The first is the case where $\alpha=n$. Here, the wave function is $\psi=e^{-ikx}e^{-i\alpha\theta}$, which is evidently single-valued when α is an integer. (It can be seen by direct differentiation that this is a solution.)

The second case is that of $\alpha=n+\frac{1}{2}$. Because $J_{(n+\frac{1}{2})}(r)$ is a closed trigonometric function, the integrals for ψ can be carried out exactly.

The result is

$$\psi=\frac{i^{\frac{1}{2}}}{\sqrt{2}}e^{-i(\frac{1}{2}\theta+r'\cos\theta)}\int_0^{[r'(1+\cos\theta)]^{\frac{1}{2}}}\exp(iz^2)dz. \qquad (23)$$

This function vanishes on the line $\theta=\pi$. It can be seen that its asymptotic behavior is the same as that of Eq. (2) with α set equal to $n+\frac{1}{2}$. In this case, the single-valuedness of ψ is evident. In general, however, the behavior of ψ is not so simple, since ψ does not become zero on the line $\theta=\pi$.

5. DISCUSSION OF SIGNIFICANCE OF RESULTS

The essential result of the previous discussion is that in quantum theory, an electron (for example) can be influenced by the potentials even if all the field regions are excluded from it. In other words, in a field-free multiply-connected region of space, the physical properties of the system still depend on the potentials.

It is true that all these effects of the potentials depend only on the gauge-invariant quantity $\oint \mathbf{A}\cdot d\mathbf{x}=\int \mathbf{H}\cdot d\mathbf{s}$, so that in reality they can be expressed in terms of the fields inside the circuit. However, according to current relativistic notions, all fields must interact only locally. And since the electrons cannot reach the regions where the fields are, we cannot interpret such effects as due to the fields themselves.

[11] See, for example, D. Bohm, *Quantum Theory* (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1951).

In classical mechanics, we recall that potentials cannot have such significance because the equation of motion involves only the field quantities themselves. For this reason, the potentials have been regarded as purely mathematical auxiliaries, while only the field quantities were thought to have a direct physical meaning.

In quantum mechanics, the essential difference is that the equations of motion of a particle are replaced by the Schrödinger equation for a wave. This Schrödinger equation is obtained from a canonical formalism, which cannot be expressed in terms of the fields alone, but which also requires the potentials. Indeed, the potentials play a role, in Schrödinger's equation, which is analogous to that of the index of refration in optics, The Lorentz force $[e\mathbf{E}+(e/c)\mathbf{v}\times\mathbf{H}]$ does not appear anywhere in the fundamental theory, but appears only as an approximation holding in the classical limit. It would therefore seem natural at this point to propose that, in quantum mechanics, the fundamental physical entities are the potentials, while the fields are derived from them by differentiations.

The main objection that could be raised against the above suggestion is grounded in the gauge invariance of the theory. In other words, if the potentials are subject to the transformation $A_\mu\rightarrow A_\mu'=A_\mu+\partial\psi/\partial x_\mu$, where ψ is a continuous scalar function, then all the known physical quantities are left unchanged. As a result, the same physical behavior is obtained from any two potentials, $A_\mu(x)$ and $A_\mu'(x)$, related by the above transformation. This means that insofar as the potentials are richer in properties than the fields, there is no way to reveal this additional richness. It was therefore concluded that the potentials cannot have any meaning, except insofar as they are used mathematically, to calculate the fields.

We have seen from the examples described in this paper that the above point of view cannot be maintained for the general case. Of course, our discussion does not bring into question the gauge invariance of the theory. But it does show that in a theory involving only local interactions (e.g., Schrödinger's or Dirac's equation, and current quantum-mechanical field theories), the potentials must, in certain cases, be considered as physically effective, even when there are no fields acting on the charged particles.

The above discussion suggests that some further development of the theory is needed. Two possible directions are clear. First, we may try to formulate a nonlocal theory in which, for example, the electron could interact with a field that was a finite distance away. Then there would be no trouble in interpreting these results, but, as is well known, there are severe difficulties in the way of doing this. Secondly, we may retain the present local theory and, instead, we may try to give a further new interpretation to the poten-

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4

tials. In other words, we are led to regard $A_\mu(x)$ as a physical variable. This means that we must be able to define the physical difference between two quantum states which differ only by gauge transformation. It will be shown in a future paper that in a system containing an undefined number of charged particles (i.e., a superposition of states of different total charge), a new Hermitian operator, essentially an angle variable, can be introduced, which is conjugate to the charge density and which may give a meaning to the gauge. Such states have actually been used in connection with recent theories of superconductivity and superfluidity[12] and we shall show their relation to this problem in more detail.

ACKNOWLEDGMENTS

We are indebted to Professor M. H. L. Pryce for many helpful discussions. We wish to thank Dr. Chambers for many discussions connected with the experimental side of the problem.

[12] See, for example, C. G. Kuper, *Advances in Physics*, edited by N. F. Mott (Taylor and Francis, Ltd., London, 1959), Vol. 8, p. 25, Sec. 3, Par. 3.

Approved For Release 2003/04/18 : CIA-RDP96-00789R003100030001-4