# AN ANALYSIS OF THE DECISION-MAKING FUNCTIONS OF A SIMULATED AIR DEFENSE DIRECTION CENTER

A. Sweetland
W. W. Haythorn
Logistics Department
The RAND Corporation

P-1988

May 11, 1960
Revised September 20, 1960

FOREWORD

Shortly after the first manual Air Defense Direction Centers became operational it became evident that air defense was not as effective as had been anticipated. Analysis strongly suggested that the cause was a "human degradation" factor: The hardware was found to be excellent; therefore, the fault must be in the way the hardware was used. As a result, a series of studies of a simulated Air Defense Center was begun in 1951 by The RAND Corporation.

With the coming of SAGE, manual air defense has disappeared and it is now possible to declassify much of the information coming from these studies. This information is offered because (1) much is germane to any system involving data processing and decision making; and (2) it is an illustration of a method of research for studying the decision processes of systems.*

---

*Credit for the design of these experiments must go to John Kennedy, then head of the Systems Research Laboratory, William Biel, Robert Chapman, and Allen Newell. Both of the present authors came to RAND months after the last of four major studies. Most of our energies were devoted to the analysis of the data yielded by the second and third experiments. This paper is the result of part of this analysis.

## DESCRIPTION OF THE SYSTEM

A manual Air Defense Direction Center is a fairly autonomous Combat Information Center having the responsibility of defending a sizable piece of geography against air attack. It is composed of three closely interacting functions:

1. The Surveillance Function. Outlying early warning radars scan the horizons and (like crude television cameras) register anything that moves. This information is displayed on a radar 'scope that is monitored by one or more humans. These humans endeavor to separate wheat from chaff and call in the significant information. This is no easy task. The definition of "significant" varies from moment to moment (the concern of this paper). Further, a radar can, and does, faithfully report all manner of "garbage:" ground clutter, storm centers, ships at sea, and (at least on one occasion) inter-urban trolleys. The picture is further complicated by weather phenomena (inversions), and relative reflectivity of various aircraft. All and all, the job is at once so difficult and so tedious that the men are rotated every hour to keep from "going ape."

2. The Identification Function. The reports from the "tellers" (scope watchers) are phoned in to the "plotters" at the Direction Center, where the track histories are plotted on a large map of the area. Most frequently this is done by grease pencil on the back of a large illuminated plexiglass status board. When a new report comes in, the plotters draw an arrowhead with the point at the new location of the track and extend the shaft of the arrow back to the previous report. They also enter other identifying information: speed, altitude, origin, etc.

This board is continuously reviewed by the Identification Group whose major activity is to match (correlate) these tracks with flight plans that have been filed by military and commercial aircraft. Once the track is identified as friendly, it is generally "scrubbed" unless there is a chance of confusing it with other tracks. Numerous techniques are used to assist in this process: plot unknowns in red and knowns in yellow, "cease-tell" until confusion is eliminated, selective plotting, etc. When several sites sumultaneously report several tracks for several minutes, the status board looks like a web made by a very confused spider.

3. The Inception Function. This is relatively separate from the first two functions. Any tracks that remain unidentified are turned over to the Interception Controller. Interception Controller scrambles an aircraft (orders it to take off) and verbally directs it via radio out to the unknown aircraft that must be identified by type and number by the pilot of the interceptor. Because scrambles are expensive and, in cases of bad weather, dangerous, every effort is made to improve the effectiveness of the Surveillance and Identification functions to avoid a scramble.

We wish to express this process in more abstract terms: The Air Defense Center, like most decision-making centers, continually samples reality attending to those elements of information germane to its charged responsibility. These elements are used to construct a model of reality, i.e., an abstraction. This abstraction is analyzed to determine meaningful ways of dealing with reality. This determination results in a response. The process is essentially first inductive and then deductive: inductive in the construction of the model, and deductive beginning with the analysis of the model.

Concurrent and intimate to this process are a number of processes related to maintaining and improving effectiveness:

1. There is a continuous updating of the model to keep pace with the changes in reality.

2. There is a continuous process of validity checking to determine if the model is an adequate abstraction of reality.

3. There is a continuous process of technique improvement: better ways of sensing, recording, abstracting, analyzing and responding.

## PURPOSES AND METHODS

With this in mind it is possible to state the intent of the experiment. The experiment was designed to vary systematically certain dimensions of the task and observe and record the effects on the processes just described. The crews were encouraged to make as many changes in procedure as they desired. The analysis described herein was focused on discovering relationships between the variable task dimensions and crew behavior.

To protect the reader we are using a somewhat unorthodox format in this discussion: Each independent variable is discussed individually (as if it were a separate experiment). It is described, its effects stated and then discussed. Most of the numbers are omitted. The reader may assume that every statement of fact has been shown to be statistically significant. Although there are some interactions, we have included them in the discussion of the main variables to avoid confusion. These are suggestively few. Indeed, we feel that, like the barking of the hound of Baskerville, there is significance in the absence. First, a brief survey of the procedures.

The crew members were selected by the Air Defense Command with the stipulation that they be of average ability. The Cowboy crew (second experiment) consisted of 39 men. Cobra (third experiment) included 40 men.

The Cobra officers, in general, were older, had higher rank and more military and air defense experience. The Cobra noncommissioned officers had a greater proportion of higher ratings, but the same amount of military and Air Defense experience as Cowboys (average, two years).

The Direction Center was comprised of surveillance, identification, and interception functions. Three early warning sites (radars) were manned by military personnel: North, West and South. West's zone overlapped with both North and South. These overlap areas were known as "cross-tell zones."

The experimental variables discussed here are:

1. Class of tracks: penetrating, local or outbound
2. Traffic load: density of air traffic
3. Crew: Cowboy vs. Cobra
4. Period: first vs. second half of each morning and afternoon
5. Experience with same load: first four days with last four days
6. Load distribution: among the three early warning sites

The experiment consisted of four sets of four days. (With a day off in between). Each day had two problems (morning and afternoon). Each problem consisted of two periods of 75 minutes each with a 40-minute nexus as the air traffic gradually changed. Thus, the design unit was the 75-minute period, of which there were 64. Each period was designed with a specific value for each of the task variables. These variables were orthogonal to all other variables. This is summarized in Table 1. This is a 2 x 3 x 2 x 2 x 2 x 3 factorial design, with two cases per cell. The statistical approach was primarily analysis of variance.

Table 1

FACTORIAL DESIGN OF THE STIMULUS

| Load | Experience | Distribution | Period | Tracks: Local | Tracks: Penetrating | Tracks: Outbound |
|---|---|---|---|---|---|---|
| 2 | 1st | Even | 1 | 2* | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| | | Uneven | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| | 2nd | Even | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| | | Uneven | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| 3 | 1st | Even | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| | | Uneven | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| | 2nd | Even | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| | | Uneven | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| 4 | 1st | Even | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| | | Uneven | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| | 2nd | Even | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |
| | | Uneven | 1 | 2 | 2 | 2 |
| | | | 2 | 2 | 2 | 2 |

*Two cases in each cell.

## RESULTS

The major results are summarized in Tables 2 and 3. The bottom row of Table 4 gives the percentage of total variance attributable to the significant sources of variance in the column. It is the square of the quantity (one minus remainder - mean - square divided by total-mean-square):

$$E = (1 - \frac{V\ rem}{V\ tot})^2$$

This term may be interpreted essentially the same way as communality is interpreted in factor analysis: as an expression of controlled variance. As shown, most of the variance is attributable to main effects and a few first-order interactions.

One further explanation is necessary: the measures. We have used four measures of track-handling behavior.

1. Number of Stimulus Tracks Carried (C)

A track was "carried" if the crew made a single response to it, no matter what the response was: an initial plot, position report, speed or altitude estimate, transfer or acceptance of tracking responsibility to or from another station, or increase or decrease in the number of aircraft carried under a particular track number. The number of tracks carried is thus a measure of what kind - and how many - of the many tracks in the stimulus the crew thought important enough to respond to.

2. Percentage of Stimulus Tracks Carried (C%)

If in a certain period the stimulus contained twenty penetrating tracks and the crew carried ten of these, C% for penetrating tracks in that period

Table 2

ANALYSIS OF VARIANCE OF MEASURES OF TRACK-HANDLING BEHAVIORS[‡]

Over-all System Performance

| Source of Variance | df | Mean Squares for Measures of Track-Handling Behavior: C | C% | TOB | TOBav |
|---|---|---|---|---|---|
| 1. Class of track | 2 | 43,777 ** | 108,456.5** | 15,706,502** | 2,671.04** |
| 2. Load | 2 | 872.5** | 7,628.5** | 70,120** | 249.04** |
| 3. Crew | 1 | 2,070 ** | 9,373.1** | 134,767** | 0.13 |
| 4. Period-within-Problem | 1 | 481 ** | 3,396.3** | 202,142** | 62.16** |
| 5. Experience (1st-2nd) | 1 | 94 ** | 413.3 | 21,459** | 13.43 |
| 6. Distribution (Even-Uneven) | 1 | 52 * | 191.7 | N.C. | 8.14 |
| 7. Class x Load | 4 | 2,492.3** | 873.7** | 351,936** | .78 |
| 8. Class x Crew | 2 | 310 ** | 1,261.7** | 31,470** | 11.45 |
| 9. Class x Period | 2 | 70 ** | 205.6 | 20,712** | 10.72 |
| 10. Class x Experience | 2 | 43.5** | 279.2 | 6,437 | 5.96 |
| 11. Class x Distribution | 2 | 25.5 | 177.1 | N.C. | .56 |
| 12. Load x Crew | 2 | 12 | 267.2 | 6,257 | 2.18 |
| 13. Load x Period | 2 | 10.5 | 40.6 | 218 | 1.00 |
| 14. Load x Experience | 2 | 24.5 | 67.6 | 9,258* | 2.25 |
| 15. Load x Distribution | 2 | 18 | 176.8 | N.C. | 3.62 |
| 16. Crew x Period | 1 | 15 | 45.9 | 1,634 | 0.02 |
| 17. Crew x Experience | 1 | 26 | 215.3 | 728 | 0.40 |
| 18. Crew x Distribution | 1 | 25 | 145.9 | N.C. | 0.72 |
| 19. Period x Experience | 1 | 5 | 59.6 | 1,911 | 9.04 |
| 20. Period x Distribution | 1 | 58 * | 580.8* | N.C. | 4.14 |
| 21. Experience x Distribution | 1 | 28 | 2.2 | N.C. | 1.56 |
| Remainder | 253 | 15.05 | 122.44 | 2,894[1] | 4.47 |
| Within Cell | 144 | 12.80 | 142.24 | 2,913[2] | 4.88 |
| Total | 287 | 372.67 | 996.58 | 119,260 | 24.70 |

[‡]Definition of measures: C - Number of tracks carried; C% - Percentage of stimulus tracks carried; TOB - Items of task oriented behavior; TOBav - Average TOB per track carried (see pp. 7 and 10).

*P >.05
**p >.01

[1]df - 261.
[2]df - 216.

Table 3

SUMMARY OF ANALYSIS OF VARIANCE SIGNIFICANCE TEST PROBABILITIES FOR OVERALL SYSTEM PERFORMANCE*

| Source of Variance | Measure of Track-Handling Behavior: C** | C%** | TOB** | TOBav** |
|---|---|---|---|---|
| 1. Class of track | .01 | .01 | .01 | .01 |
| 2. Load | .01 | .01 | .01 | .01 |
| 3. Crew | .01 | .01 | .01 | --- |
| 4. Period-within-Problem | .01 | .01 | .01 | .01 |
| 5. Experience (1st-2nd) | .01 | --- | .01 | --- |
| 6. Distribution (Even-Uneven) | .05 | --- | N.C. | --- |
| 7. Class x Load | .01 | .01 | .01 | --- |
| 8. Class x Crew | .01 | .01 | .01 | --- |
| 9. Class x Period | .01 | --- | .01 | --- |
| 10. Class x Experience | .05 | --- | --- | --- |
| 11. Class x Distribution | --- | --- | N.C. | --- |
| 12. Load x Crew | --- | --- | --- | --- |
| 13. Load x Period | --- | --- | --- | --- |
| 14. Load x Experience | --- | --- | .05 | --- |
| 15. Load x Distribution | --- | --- | N.C. | --- |
| 16. Crew x Period | --- | --- | --- | --- |
| 17. Crew x Experience | --- | --- | --- | --- |
| 18. Crew x Distribution | --- | --- | N.C. | --- |
| 19. Period x Experience | --- | --- | --- | --- |
| 20. Period x Distribution | .05 | .05 | N.C. | --- |
| 21. Experience x Distribution | --- | --- | N.C. | --- |
| Percentage of total variance attributable to significant effects | 92 | 76 | 95 | 69 |

*N.C. indicates the mean square was not computed because of its obvious lack of significance. The dashes indicate non-significant effects.

**Definition of measures: C - Number of tracks carried; C% - Percentage of stimulus tracks carried; TOB - Items of task oriented behavior; TOBav - Average TOB per track carried (see pp. 7 and 10).

is 50 per cent. Using the C% value has a number of advantages, particularly in comparing the crew's responses to different classes of tracks. On the other hand, the use of this percentage obscures changes in actual numbers (the C value). Fifty per cent of twenty penetrating tracks is not equivalent to 50 per cent of sixty penetrating tracks in considering the activity level of the crew.

3. Items of Task Oriented Behavior (TOB)

Each item of track-handling information was counted once in tallying the value of TOB. Although this method is seemingly arbitrary, these are the units of information a crew uses in constructing its picture of a flight. Counting them gives a rough measure of the amount of crew activity.

4. Average TOB per Track Carried (TOBav)

The average number of responses (or items of task information) per track is a measure of surveillance activity level per unit track. Changes in TOBav yield clues about what is causing a crew to change its procedures.

## DISCUSSION

1. Classes of Tracks

This factor accounted for most of the variance. The stimulus was built of three classes of tracks: (1) penetrating tracks, those originating outside of the system and seen coming towards the system; (2) outbound, those going away from the system; (3) local, those remaining within the system. From the view of the participants these three classes of tracks were of differing importance. Any penetrating aircraft may be hostile (from time to time both mass raids and "sneaks" - single hostiles - were included). In contrast any track leaving the system probably is not hostile. And any originating in the system obviously is not dangerous. (The chance that an aircraft taking off from a local airport is an armed Russian bomber is, seemingly, remote.)

The differences of response between these two classes (potentially dangerous, therefore important; probably innocuous, therefore unimportant) were the numbers (and proportions) carried. Important tracks were carried, unimportant tracks dropped. This is discussed more meaningfully following the discussion of load. At the present we wish to point out that this differentiation is but one of a number of orderings along an "importance" continuum made by the crews. For example: tracks penetrating from dangerous directions (directions associated with previous hostile attacks) were carried far more often than tracks from non-dangerous directions.

In general, crews maintained "important" tracks and eliminated unimportant tracks. This process was called "filtering". Our conjecture is that the stimuli were ordered along a subjective continuum varying from "no response is necessary" to "immediate response is mandatory" and that the other variables helped determine this ordering. This conjecture is based on the fact that most of the interactions (Tables 2 and 3) were with the class of track variable.

2. Load

The tracks were computed, the original information being on decks of IBM cards. A "deck" consisted of 36 tracks: 14 background ("noise") tracks, 12 penetrating, and 10 outbound. One deck was used in Load 1, two in Load 2, etc. Figure 1 shows the effect of load changes. (Figure 1 includes Loads 1 and 5. These two loads are not included in the statistical computations as they are not truly orthogonal, They were used chiefly as cross validation checks in this study.)

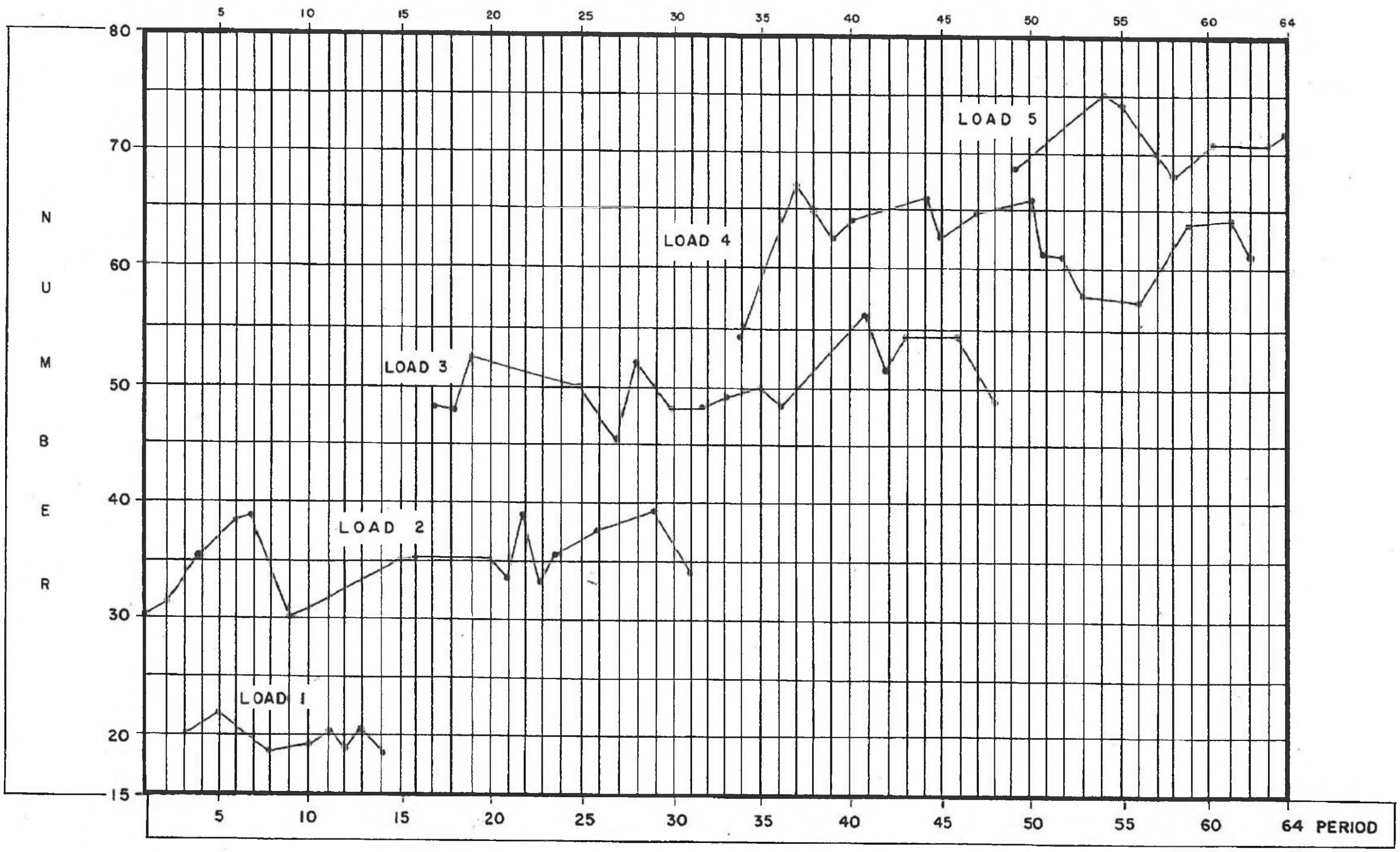



FIGURE · NUMBER OF PENETRATING TRACKS CARRIED IN EACH PERIOD

At first glance (Fig. 1) the crews behaved like a single-pole, 5-position switch: Load 1 yielded Load 1 behavior, Load 2 yielded Load 2 behavior, etc. This does not, however, tell all of it. The proportion of penetrating tracks from dangerous directions (and therefore potentially hostile) remained essentially the same (90 - 95 per cent) regardless of the load increase. But there was a marked decrease in the proportion of unimportant tracks carried (interaction of class and load).

The increase in load provoked the elimination of non-essential behavior. Load increases finally caused a pruning of almost all behavior not critical to defending the area. As will be shown, the crews also (as load went up) tended to carry tracks for shorter and shorter times, and also with fewer and fewer reports.

Thus the increase in load causes a modification of the model of reality made by the crews. At the beginning (under low loads) the crews tended to include everything (penetrating, noise, outbound) in the model. As the loads increased they dropped the non-essentials, still maintaining an accurate replication of the essentials. This one to one relationship between important reality and model remained. Rather than yield, they made improvements in their processing techniques (quicker decisions, fewer reports). Our feeling is that an excellent way to measure the importance of a phenomenon to a system is to determine the accuracy of the organization's model of reality with respect to the phenomenon (the greater the accuracy the greater the importance).

3. Crew Differences

Although there were differences between the two crews, we are inclined to minimize the significance. The chief difference was in the number of tracks carried (Cobra carried more). Closer inspection of the data showed that this was chiefly due to the number of outbound tracks carried. The least difference between the two was in the handling of the penetrating tracks: Both crews maintained close, accurate contact with the stimulus on these critical tracks.

One crew (Cobra) was more authoritarian, the other (Cowboy) more democratic in approach (in the Frankel-Brunswick sense). We found some suggestions that the younger and "more democratic" Cowboy was more willing to try novel procedures which may have accounted for the fact that Cowboy drastically reduced the number of track confusions the last four days. (These increased for Cobra.) But this had little effect on overall effectiveness.

In short, when it came to accurate handling of critical tracks, the crews performed similarly, and their behavior was essentially stimulus bound. There were some differences in the way they handled the non-essentials: the more conservative Cobra seemed less willing to prune out non-essentials and, as a result, worked harder in maintaining the same level of effectiveness. There was a slight indication that the reluctance to prune might have proved costly had the experiment gone to Load 6. It is more likely, however, that Load 6 would have provoked some hearty pruning by Cobra (i.e., as evidenced by the load by class interactions).

We cannot resist pointing out one fact: the results give no support to the champions of either the authoritarian or democratic approach. As a matter of fact, the data strongly suggest that when it comes to critical situations the question is academic: both respond to the needs of the situation, the etiquette of procedures being shelved until the crisis passes.

4. <u>Period Effects</u>

One problem was run each half-day (forenoon and afternoon) and each problem consisted of two periods (the experimental unit). Comparison of the first and second periods showed greater activity (more tasks carried) during the first period. The crews were unaware of the period differences: There was a transition between them as the air traffic gradually unfolded from one situation to the other with no discontinuity.

Although a number of conjectures were offered (fatigue, boredom, motivational changes), none of the data offered substantiation. It was not until we took a more microscopic look at the data that we discovered what was happening.

At the beginning of each problem the crew was confronted with a large number of tracks that had to be sorted out, so they carried everything until it was identified (or recognized as being non-threatening). Once they had the continuity completed, the "plot everything" behavior disappeared. Incidentally, further detailed analysis showed no evidence of a "fatigue factor." The last half hour of every day, until the close of the experiment, found the crews still actively plotting tracks at the same pace.

We wish to elaborate on this Period Effect. At the time, we closed our books satisfied that we had accounted for a factor peculiar to a specific simulation. However, we have noticed this same effect in a number of subsequent simulations. The effect occurs whenever a novel element in introduced into the system. At this point there is an "initial burst" activity associated with making an analysis of the new phenomenon. The characteristics of this activity are different from the processing activities: the initial burst activity is essentially diagnostic. The new phenomenon is analyzed and the results of this analysis determine the treatment (processing) to be used.

There are two reasons, in experimental work, to bear in mind the dichotomy between diagnosis and treatment. First, subjects are reluctant to proceed until they have fully analyzed the job. Accordingly, time to analyze and digest must be provided. Second, the presence of the diagnostic activity in the data contaminates the picture of the processing aspects of the operation. Consequently, it is expedient to keep these separate (if possible). An illustration of the importance of separating the two is given in the discussion of the results of the microscopic analysis.

5. Experience with the Same Load

Since loads 2, 3, and 4 were each used in two sets, eight times as a high load in one set, then eight times as a low load in the following set, it was possible to measure the change resulting from experience (learning?) by comparing the first eight with the last eight.

These comparisons showed tiny (but statistically significant) reductions of the number of tracks carried. A few interactions with other variables showed in a more detailed analysis, but these were erratic (occurring mostly in the data from the southern early warning station).

In general we were more impressed by the absence of evidence of learning, than the presence. We are inclined to attribute this absence to the relatively short time the crews were in the experiments as compared with their average air defense experience (two years). The microanalysis will discuss this learning factor in another dimension.

6. Distribution

This was included as an "even-uneven" variable. During "even" periods the load was distributed fairly evenly between the North and West early warning stations. During "uneven" periods West's load was increased while North's load was decreased. This was done to investigate "load balancing": it would be possible for North to carry all tracks in the overlap zone, when West's load was inordinate.

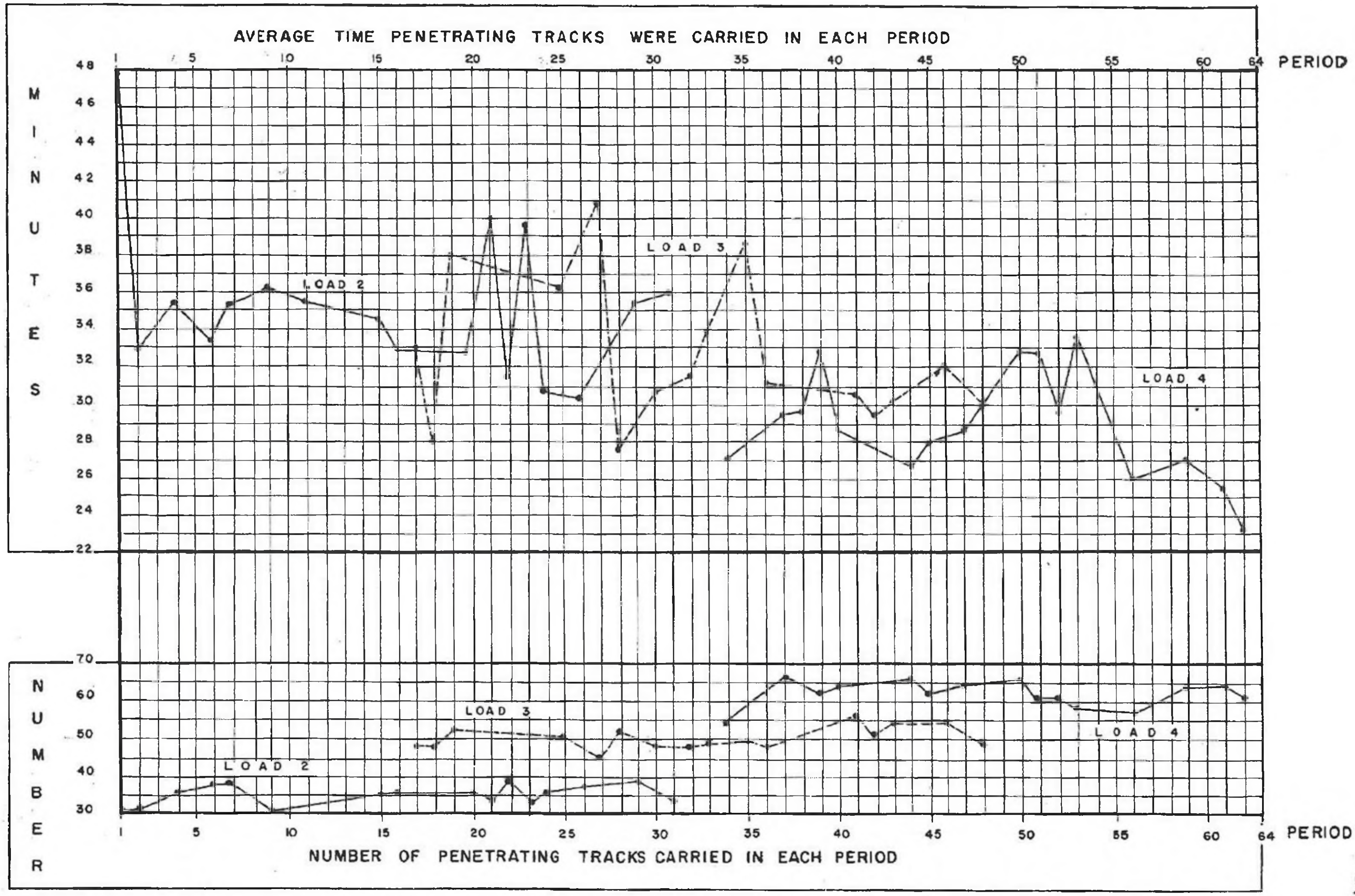
The data showed no definite evidence of load-balancing. Nor did any more detailed analysis show load-balancing. We have a strong suspicion that the statistical significance we obtained may be a false positive: several hundred F-tests were made during the analysis. This number insures the inclusion of some "significant" results that are really random variations. Our practice has been not to attach importance to an observation unless several different measures all point toward the same interpretation. This latter did not occur with Distribution.

## 7. Further Analysis

Thus far the discussion has been concerned with what was found in comparing aggregates of 75-minute periods. A number of the results suggested that a more detailed analysis was in order. Accordingly, this was done. Most of the following findings are based on data taken in 5-minute aggregates. Only one crew (Cobra) was used in this analysis (budget limitations preventing the detailed analysis of both).

The first question concerned the average time the various classes of tracks were carried under the several experimental conditions. It was found that as the experiment progressed, there was a decrease in the average time tracks were carried. The careful phrasing of the previous sentence is necessary: the design resulted in a confounding of load and experience. To determine which (if either) was predominant, the data were plotted against time (see Fig. 2). The general slope (without marked separations by load) shows that, as an analysis of variance suggested, experience was the predominating factor.

A second analysis concerned the communication rate. The classes of tracks were separated and the numbers of responses were divided by the total




COMPARISON OF TRACKS CARRIED AND AVERAGE TIME CARRIED BY COBRA CREW

- Figure 2 -

time track-continuity was maintained. The result is a response-per-track-minute, in effect a "chatter rate." These results were almost perfect negative correlations with the previous TOB data: the chatter rate apparently increased with load and experience.

The close relationship between TOB and C is seen in the following: During any given five-minute period the Cobra crew carried from two to seventeen tracks. This required the passing of a varying amount of track-handling information. To determine the relationship between TOB and C, the average amount of TOB per track was plotted against the number of tracks being carried (both measured in five-minute units).

The resultant curve, unusually regular for psychological data, was fitted as a log function. As can be seen in Fig. 3, the data fell closely around a line fitted by the equation:

$$TOB = 20.25 (\log C) - 2.56$$

in which "TOB" is the mean number of Task-Oriented-Behaviors per five minutes and "C" is the associated number of tracks being carried.

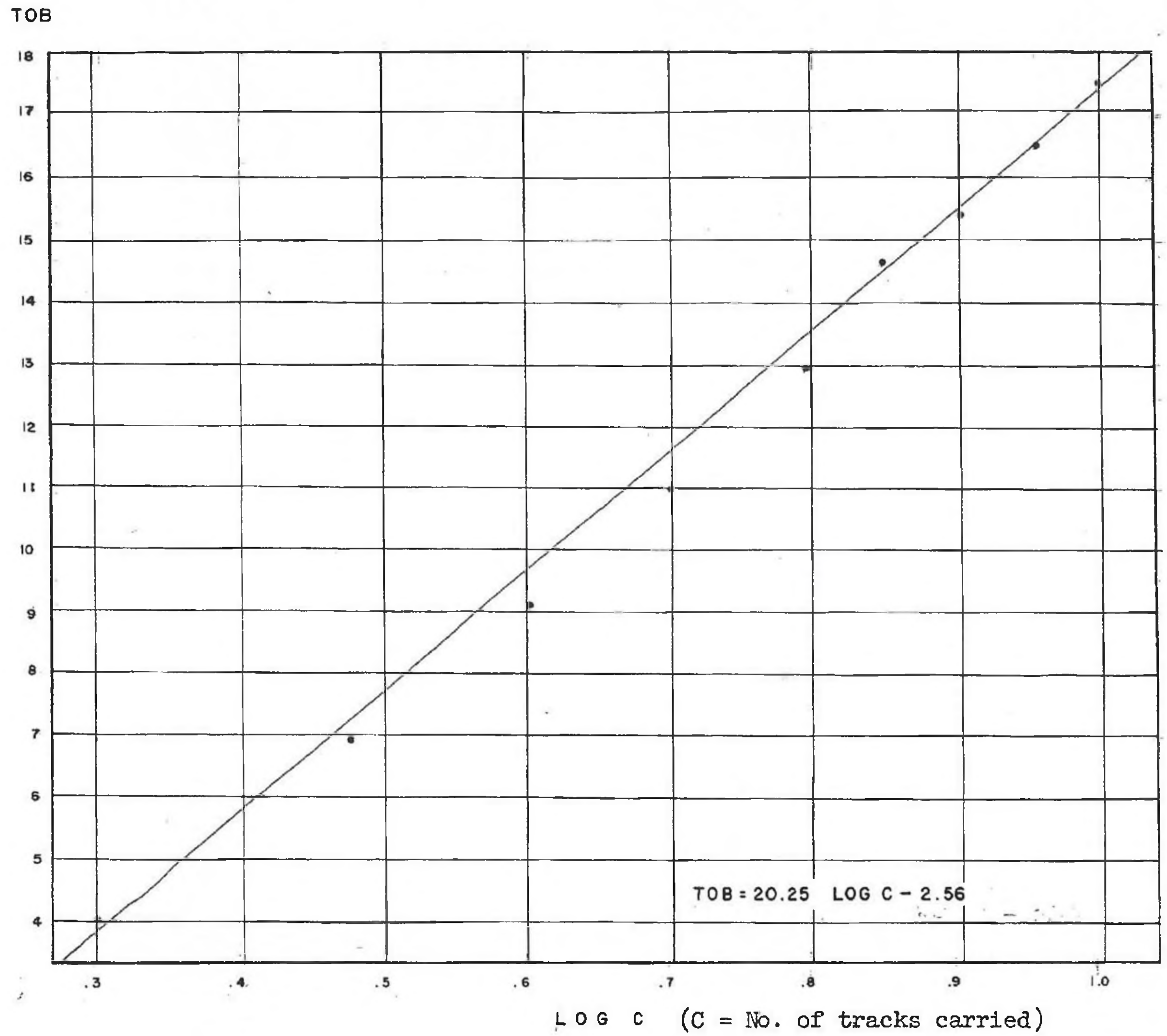
This is interesting in view of the fact that the equation of best fit is identical to Fechner's law relating sensation to stimulus strength:

$$S = K (\log R)$$

in which S is equal to sensory magnitude, R is the stimulus measured from absolute threshold, and "K" is a constant. Thus the question arises as to whether it's possible to think of TOB as being a measure of sensory magnitude with C as its stimulus.

Further study showed that when a track is initiated the crews gave:

1. an initial position report
2. an altitude estimation




RELATIONSHIP BETWEEN TASK ORIENTED BEHAVIOR AND NUMBER OF TRACKS CARRIED

- Figure 3 -

3. a speed report

4. a track number

5. a second position report (establishing direction)

Following this the track was carried by periodic position reports until it was scrubbed. Thus, a tracking activity may be divided into two fairly distinct activies: 1) initial burst and 2) carrying.

Our hypothesis that the initial burst activity caused the apparent increase in activity was tested by subtracting out the initial burst. It was then seen that the "chatter rate" remained constant throughout the experiment (i.e., that it did not change with load, experience, time, etc.).

The data also suggested looking at the number of simultaneous tracks carried. The results showed only a slight increase throughout the experiment. What was happening was that the crew put out essentially the same amount of effort throughout the experiment, but they were still able to handle more tracks by maintaining them under surveillance for shorter and shorter times. (The number of simultaneous tracks carried remained essentially constant: six-to-seven per plotter-teller combination. This was confirmed by analyzing the relationship between initiations and scrubs. It was found that 80 per cent of initiations occurred within two minutes following a scrub.)

With all these factors in mind it is possible to formulate a picture of what went on.

1. From the beginning the crews said, in effect: "We will carry, at any point in time, six aircraft per plotter-teller combination. Further, we will report on each aircraft once every three minutes."

2. They also said, in effect: "We will pick the six most important aircraft to maintain."

There were occasional exceptions: If fewer than six penetrating aircraft were present, they were willing to carry more than six (apparently because local and outbound tracks were easier to carry). Or, if the situation demanded, they would carry more than six criticals. The largest was 17 in a five-minute period. This happened once out of 61,000 five-minute periods.

It appears that most of these results can be explained by a series of simple, dichotomous (yes-no) decisions. This is given in Fig. 4.

Thus, there seem to be at least three major elements in a large decision-making system:

A. Reality Modeling. When reality is too complex to handle directly, a model is made of this reality. The major characteristics of this model reflect the charged responsibility, in that the model will consist chiefly of those elements that must be controlled in order to meet the objectives of the system. The model construction is essentially an inductive process. It yeilds a simplified picture of reality. It's intent is of a diagnostic nature.

B. Model Analysis. The model is analyzed and the results of this analysis determine the responses of the system. These are carried out. This is essentially deductive in nature.

C. Energy Level. Prevading all this appears to be a norm that declares how hard the group or system will (or can) work to obtain the objectives. This tends to remain constant. If it happens that reality intrudes a heavier load than this normative rate permits, the least

Stimulus Track

a. Detected? (seen) ⟶ no — no report
↓
Yes
↓
b. Is it an established track? ⟶ yes — report it
↓
No
↓
c. It it potentially critical? ⟶ no — forget it
↓
Yes
↓
d. Are we already carrying our limit? ⟶ no — report it
↓
Yes
↓
e. Can we drop something less important? ⟶ no — delay action
↓
Yes
↓
f. Drop old track and report new.

Fig. 4 -- DICHOTOMOUS DECISION MODEL

important activities are pruned, techniques are made more efficient, or better techniques are devised. In short, everything reasonable is done to keep from exceeding the system's norm. This may be due to physiological, system-structural, social, or other reasons.

One final observation: Interactions were significant by their absence. The only interactions that existed were first-order interactions, and most of these were the most obvious (i.e., interactions with class of tracks). Indeed, in doing the statistical computations, it was quickly discovered expedient to terminate computations with the first-order interactions by testing against the within groups. Almost inevitably the remaining variance was less than the within groups.

Our thesis is that, without formal training, the simplest first-order interaction is difficult for the average human being to comprehend even when the data are before him. Expecting people to show differential methods of operation in a situation as complex and amorphous as this would be expecting too much. Perhaps with months of intense experience interactions might have appeared (i.e., as they do in a thoroughly explored situation such as professional football).

It seems to us, however, that a more expedient approach (as opposed to accumulating experience) would be for systems analysts to investigate the area and, analytically, determine which areas would yield a profit if the responses were based on interactive relationships. These should be stated as decision rules: "Do X only if A is black and B is large (if A is not black or B is not large, do Y)."

## SUMMARY AND CONCLUSIONS

Two crews from the Air Defense Command were charged with the responsibility of defending the United States against hostile attack and then exposed to high stress situations in a simulated manual Air Defense Direction Center. Their responses were analyzed to determine what, if any, characteristic response patterns emerged. The following were suggested by the analysis of the data:

1. When reality is too complex to deal with first-hand, a man-machine system will construct a model of reality. This model is continuously modified by the system (to take into account changes in reality and changes in proficiency) but the modifications always reflect the charged responsibility.

2. The model of reality is then evaluated to determine the best of several possible responses. Responses are selected in terms of which most effectively meets the charged responsibility.

3. A system, apparently unconsciously, operates at a level of effort that remains relatively constant for the job at hand and will exceed this level rarely and for short intervals only.

4. A system apparently orders reality on a subjective list ranging from "immediate response is mandatory" to "no response is necessary" and then responds reading from the top of the list to the point of the system level of effort.

5. The major effect of stressing the system is to provoke a more careful culling of the list rather than increase the level of effort.