AD-A099 908 RAND CORP SANTA MONICA CA F/G 5/10
SUBJECTIVE MEASUREMENT OF TACTICAL AIR COMMAND AND CONTROL. VOL--ETC(U)
MAR 81 C T VEIT, B J ROSE, M D CALLERO F49620-77-C-0023
UNCLASSIFIED RAND/N-1671/3-AF NL

1 of 1
AD
A099908

END
DATE
FILMED
6 81
DTIC

BS

1- AD-A099145
2-

A RAND NOTE

SUBJECTIVE MEASUREMENT OF TACTICAL AIR COMMAND AND CONTROL--VOL. III: PRELIMINARY INVESTIG OF ENEMY INFORMATION COMPONENTS

Clairic[illegible]. Veit, Barbara J. Rose,
Monti [illegible] Callero

March 1981

N-1671/3-AF

Prepared For The United States Air Force

Rand
SANTA MONICA, CA. 90406

Published by The Rand Corporation

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER N-1671/3-AF	2. GOVT ACCESSION NO. AD-A099 908	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Subjective Measurement of Tactical Air Command and Control--Vol III: Preliminary Investigation of Enemy Information Components		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Clairice T. Veit, Barbara J. Rose, Monti D. Callero		8. CONTRACT OR GRANT NUMBER(s) F49620-77-C-0023
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Rand Corporation 1700 Main Street Santa Monica, CA. 90406		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Requirements, Programs & Studies Group (AF/RDQM) Ofc, DCS/F&D and Acquisition Hg USAF, Washington, DC 20330		12. REPORT DATE March 1981
		13. NUMBER OF PAGES 50
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release: Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

No Restrictions

18 SUPPLEMENTARY NOTES

19 KEY WORDS (Continue on reverse side if necessary and identify by block number)

Tactical Command and Control
Evaluation
Measurement
Cognition
Hypotheses

20 ABSTRACT (Continue on reverse side if necessary and identify by block number)

See Reverse Side

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Three different experiments employed subjective transfer function techniques in analyses of the enemy information section of the tactical air command and control and force employment representation described in Volume II of this series. In all three experiments, Air Force professionals judged questionnaire items that described characteristics of enemy information against a Korean-like land battle scenario. In two experiments, judgements were of the value of the enemy information characteristics for knowing the enemy's capabilities to conduct ground operations against friendly forces. Results indicated that Precision, Amount and Currency were the appropriate enemy information characteristics to include in the representation, and a range model was the appropriate subjective transfer function to explain the observed divergent interactions among the characteristics; when one characteristic was poor, the other characteristics had less of an effect on the value judgment. In the third experiment, respondents judged the Currency of the enemy information given the Frequency with which enemy second echelon forces were observed and the Time to get that information to the command and control system. A Divergent interaction was also observed in the judgments of these characteristics. The third experiment demonstrated how a representation can be extended to include characteristics (Frequency and Time) that might be easier for the decisionmaker to alter, by employing an existing characteristic (Currency) as a dependent variable.

(12)

A RAND NOTE

6

SUBJECTIVE MEASUREMENT OF TACTICAL AIR COMMAND AND CONTROL, ~~VOL.~~ III. PRELIMINARY INVESTIGATION OF ENEMY INFORMATION COMPONENTS

Clarice T. Veit, Barbara J. Rose
Monti D. Callero

March 1981

(14) RAND / N-1671/3-AF

15

Prepared For The United States Air Force

Rand
SANTA MONICA, CA. 90406

PREFACE

This is the third in a series of Rand Notes using subjective measurement to evaluate tactical air command and control and demonstrating a newly formulated subjective measurement approach, the Subjective Transfer Function approach for evaluating complex systems. The evaluation focuses on the effect of having more or less information about the enemy available to the tactical air control system.

Volume I (N-1671/1-AF) provides a background to tactical air command and control evaluation, discusses subjective measurement as an evaluation technique, summarizes the subjective transfer function approach, and describes the evaluation problem and its conflict environment. Volume II (N-1671/2-AF) describes a hierarchical representation of tactical air command and control as it is related to the employment of tactical air forces in affecting the outcome of a land battle. This representation forms the framework for initial subjective measurement experiments, and its evolution and development during the course of the evaluation are an integral part of the subjective transfer function approach.

This Note, Volume III, presents results from a set of three preliminary experiments that applied the subjective transfer function approach to a portion of the representation described in Volume II. The objective of these experiments was to test hypotheses about the (1) "appropriate" components for describing Enemy Information in the representation, and (2) "appropriate" transfer functions to explain observed effects of component descriptions on judgments of their value for knowing the enemy's capability to conduct ground operations against friendly forces.

The work was performed under the Project AIR FORCE research project "Tactical Air Command and Control." It should interest those Air Force officers concerned about the effectiveness of tactical command and control, the benefits of suggested hardware or software improvements, and criteria for measuring both the above.

SUMMARY

Two experiments were conducted to investigate the value of different characteristics of Enemy Information for controlling attacks on fixed targets in a Korean-based land battle. Two separate sets of characteristics (components) were investigated for possible inclusion in one portion of an hypothesized tactical air command and control and force employment hierarchical representation. Components hypothesized to affect the Currency of Enemy Information were investigated in a third experiment. In all three experiments, the same six Air Force professionals served as respondents. Algebraic modeling methods that have been incorporated in the subjective transfer function approach to complex system analysis (Veit and Callero, 1981) were used in these experiments to assess the effects of hypothesized components on judged outcomes.

The first two experiments used an experimental design called "RECIPE" that permits tests of hypothesized models that specify how the components affect judgments. In these experiments, Air Force professionals judged the value of different sets of Enemy Information characteristics for knowing the enemy's capabilities to conduct ground operations against friendly forces. In one experiment, the sets of Enemy Information characteristics (questionnaire items) described the proportion of enemy objects or actions of military value for which data were available (Amount), how old those data were in hours (Currency), and the degree of precision in the data on fixed targets and enemy vehicles (Precision). In the second experiment, Enemy Information was

described in terms of its Precision, Observation Frequency (the frequency with which the enemy second echelon area was observed), and Reporting Time Interval (the time elapsed from observation of an enemy event--an enemy object or action of military value--to receipt of that information by the the control function).

For both experiments, judgments exhibited systematic interaction effects among all of the characteristics of Enemy Information on judgments. In the first experiment, these interactions diverged toward higher valued information for all subjects, indicating that when one kind of information was bad (e.g., Currency was poor), the other type (e.g., Precision) had less of an effect on the value judgment. The data were consistent with a range model that predicts that subjective judgments are related not only to the relative average of the values placed on the characteristics of Enemy Information contained in an item, but also to the subjective range of values placed on those characteristics. From the model, it is possible to assess respondents' value tradeoffs that affected their judgments.

In the second experiment, individual differences were found in the form that the interaction took. Divergent interactions were found for some respondents and convergent interactions for other respondents. A range model gave a good account of these data and indicated that the individual differences lay in the weights placed on the Enemy Information characteristics. The results of these two experiments led to the conclusion that the components Precision, Amount, and Currency were the best candidates for inclusion in the tactical air command and control and force employment representation because they resulted in one set of model parameters for this portion of the hierarchical representation.

In the third experiment, judgments of Currency were investigated as a function of levels of Observation Frequency and Reporting Time Interval (as defined in the second experiment). The "scale-free" design used in this study made it possible to examine the subjective tradeoffs between levels of these two components in judgments of the Currency of Enemy Information. The scale-free design revealed a systematic divergent interaction between Observation Frequency and Reporting Time Interval in judgments of Currency.

This experiment demonstrated how hierarchical representations can be extended to levels that can be defined along dimensions that are measurable on a physical continuum (e.g., time, quantity) and thus are of more practical significance to the decisionmaker. Differences and advantages of using the scale-free and RECIPE designs in complex system analyses are discussed.

CONTENTS

Section

FIGURES

TABLES

I. INTRODUCTION

The subjective transfer function approach to complex system analysis (Veit and Callero, 1981) was developed as an analytic tool to test numerous cause and effect hypotheses embedded in the functioning of a complex system. The basic ideas of the approach are:

1. A complex system representation (i.e., stipulation of the components that define the system and their interrelationships) is generated from a series of testable hypotheses concerning causes and effects within the system that ultimately affect overall system outcomes.
2. Tests of these hypotheses provide a basis for accepting or rejecting
   a. The components as appropriate for describing the system,
   b. Hypothesized subjective transfer functions that specify the causal relationships among the components and the system outcomes.

Volume I (Callero, Naslund, and Veit, 1981a) of this series of notes defined the evaluation problem, explained the subjective transfer function approach, and described an exemplary conflict environment. The first step in applying the subjective transfer function approach is to construct an initial representation of the problem domain.

Volume II (Callero, Naslund, and Veit, 1981b) presented an initial system representation for the specific problem of interest. This representation reflects the initial hypotheses of Air Force professionals about the important components of the domain and how they are interrelated.

In this note we apply general ideas and methodology of the subjective transfer function approach to one portion of the hypothesized tactical air command and control and force employment system representation presented in Volume II. That representation is shown in Fig. 1.

The numbers 1 and 20 through 26 in Fig. 1 denote the component groups (referred to as experimental units) that correspond to initial system hypotheses. These hypotheses are not formally stated in the representation shown in Fig. 1. Rather, the representation simply depicts which components are hypothesized to directly affect other components in the system. All components are hypothesized to either directly or indirectly affect the land battle.

In the subjective transfer function approach, these hypotheses are formalized by postulating algebraic judgment models (referred to as transfer functions for reasons described in Veit and Callero, 1981) that specify the effects of the components on judged outcomes. Hypothesized transfer functions are tested within each experimental unit. The goal is to diagnose an appropriate transfer function for each experimental unit in the representation. When an appropriate transfer function is found (i.e., a tested model accounts for the judgment data), subjective scale values associated with the stimuli (component descriptions) and responses (judged outcomes) are derived from the model. The validity of the scale values rests with the tested and verified validity of the

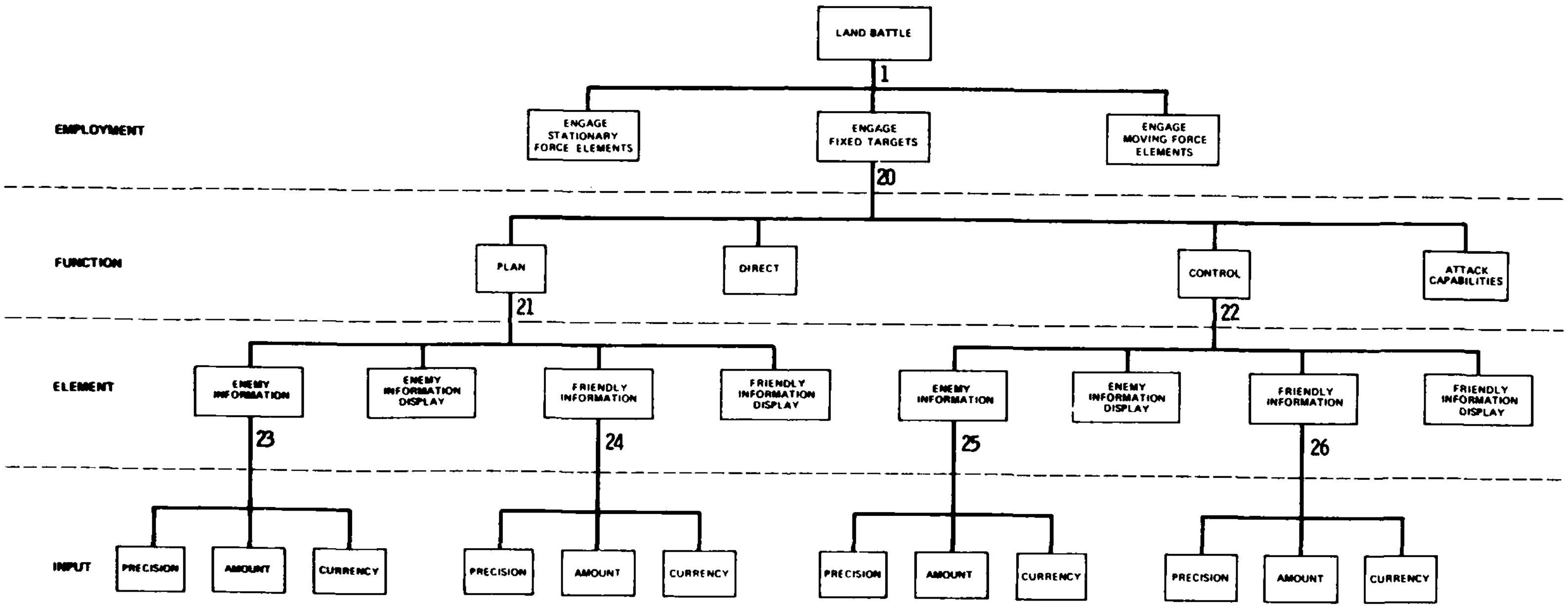

Fig. 1 — Hypothesized representation of tactical air command and control and force employment system

model.[1] When transfer functions have been determined for all experimental units, they can be used to assess how different inputs to the system alter outcomes within the system as well as the land battle (see Veit and Callero, 1981, Fig. 8).

The present research demonstrates experimental designs and general methodology to be used for studying hypothesized system components and developing transfer functions in experimental units throughout the command and control and force employment representation shown in Fig. 1.

Three pilot experiments focused on components hypothesized to describe Enemy Information (experimental Unit 25 in Fig. 1). Air Force professionals hypothesized two alternative sets of Enemy Information components. These are shown in Fig. 2. Both sets of components describe characteristics of the Enemy Information available for controlling attacks on fixed targets.

Two experiments were conducted to determine the appropriateness of each set of components for defining Enemy Information and the model (transfer function) that explained the effects of the components on judgments. A third experiment was designed to further investigate the Currency component (Panel A, Fig. 2).

THE MODELS

Determining an appropriate model to describe judgments is crucial to understanding how components of a system can be changed so as to

---

[1] The idea that scale values are derived from an appropriate model is the major characteristic of the algebraic modeling approach to measurement. See Anderson (1970, 1974a, 1974b); Birnbaum (1974); Birnbaum and Stegner (1979, 1980); Birnbaum and Veit (1974a, 1974b); Krantz, Luce, Suppes, and Tversky (1971); Veit (1978). The subjective transfer function approach applies these basic ideas to complex system analyses.

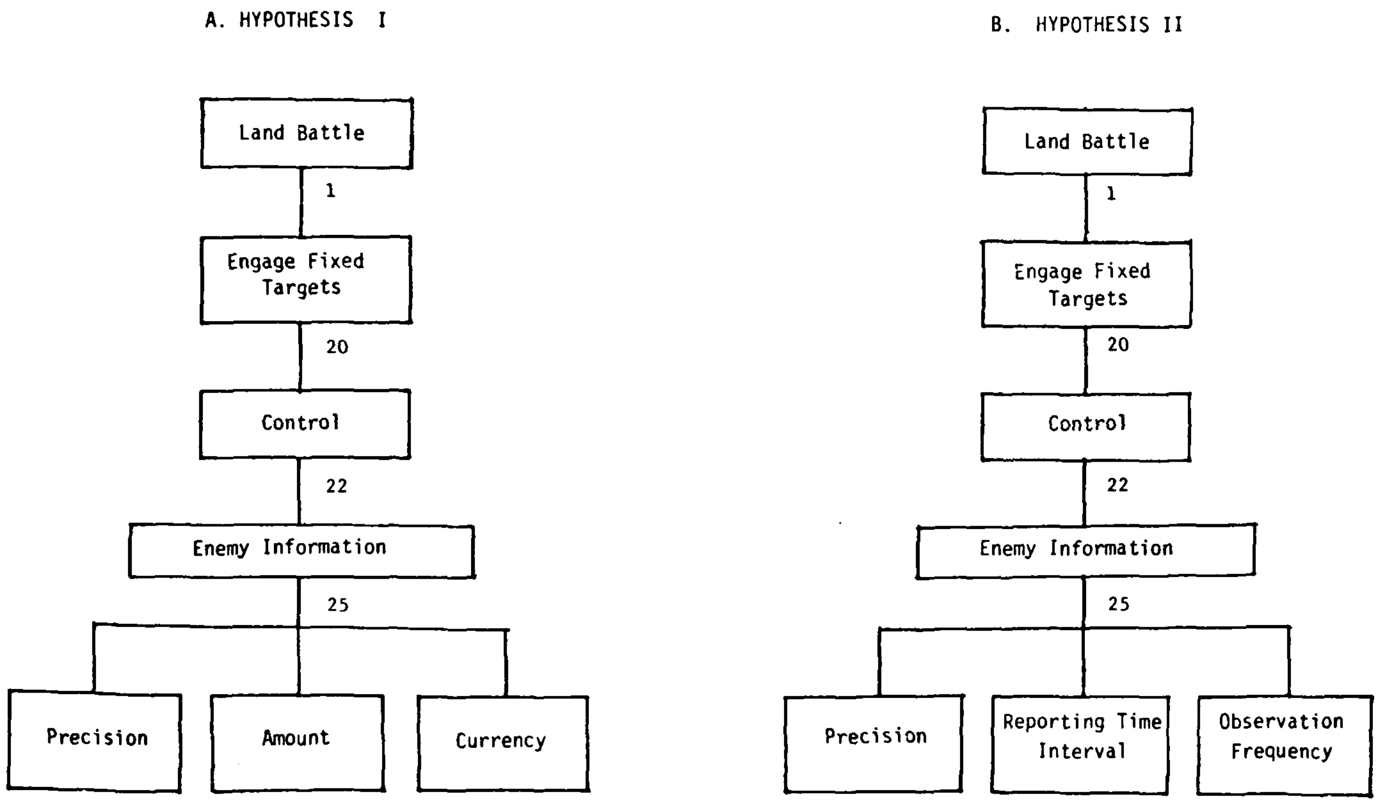

Fig. 2—Two sets of components hypothesized to affect Enemy Information
(see experimental Unit 25 in Fig. 1)

change judgments of important system outcomes. The model (transfer function) specifies how the components affect judged outcomes.

The models described below were selected initially as possible explanations of judgments concerning Enemy Information because of their success in other domains. For ease in understanding the models' predictions, the first set of components shown in Panel A, Fig. 2--Precision (P), Amount (A), and Currency (C)--will be used as the independent variables to describe the models. Imagine that the dependent variable (the response) the models are hypothesized to explain comes from the following situation that is similar to the actual experimental situation. An Air Force professional is presented with a number of questionnaire items. Each item describes the level (e.g., a particular quality or quantity) of the Precision, Amount, and Currency of the Enemy Information coming into the command and control system. The Air Force professional judges each item in terms of the value of the incoming Enemy Information for knowing the enemy's capabilities to conduct ground operations against friendly forces. Each model described below makes a different prediction about what these judgments should be. The goal is to find the model that explains the responses.

Relative-Weight Averaging Model

For three components, Precision (P), Amount (A), and Currency (C), the relative-weight averaging model can be written

$$R_{P_i A_j C_k} = J\left[\frac{w_0 s_0 + w_P s_{P_i} + w_A s_{A_j} + w_C s_{C_k}}{w_0 + w_P + w_A + w_C}\right], \qquad (1)$$

where $R_{P_i A_j C_k}$ represents the observed response to the ith level of Precision combined with the jth level of Amount and kth level of Currency; J is the judgment function that transforms the combined subjective response into an observed response, $w_0 s_0$ are the weight and scale value, respectively, of the initial impression (i.e., what the judgment would be in the absence of specific information), w_P, w_A, and w_C are the subjective weights associated with the three components, and s_{P_i}, s_{A_j}, s_{C_k} are the scale values associated with the ith, jth, and kth descriptions of Precision, Amount, and Currency, respectively. The relative-weight averaging model is a special case of a simple additive model (Anderson, 1974a). When observed responses are assumed to be linearly related to underlying subjective responses (i.e., J in Eq. 1 is linear), the major prediction for all additive models is independence (no interaction) among the components presented for judgment; thus, the effect of one of the components (e.g., Currency) should be independent of the level of the other components (Precision and Amount) presented for judgment.

This independence prediction can be seen graphically in Fig. 3. When mean responses are plotted as a function of the levels of one component (Currency) with a separate curve for each level of another component (Precision), the curves should be parallel. The curves in Fig. 3 follow this prediction exactly. The effect of Currency can be seen from the slopes of the curves. The slopes are all the same, independent of the level of Precision of the Enemy Information. Similarly, the effect of Precision is evident from the separations between the curves; for any two given curves, the vertical separation between them remains the same across all levels of Currency. If the data for the experiments performed here exhibited the parallelism shown in Fig. 3 for all three possible component pairs, the relative-weight averaging model would be supported as an explanation of the data. (We prefer graphic analyses of data (e.g., Fig. 3) to statistical analyses because they allow assessment of magnitude and direction of deviations from model predictions.)

Support for this model has been found for a number of judgment domains (Anderson, 1971; Birnbaum, 1976; Birnbaum, Wong, and Wong, 1976).

Range Model

The range model predicts that subjective responses to stimulus combinations are related not only to the relative average of the stimulus values but also to the subjective range of the stimuli presented for judgment on a particular trial. For Precision, Amount, and Currency, the response to the combination of the ith level of

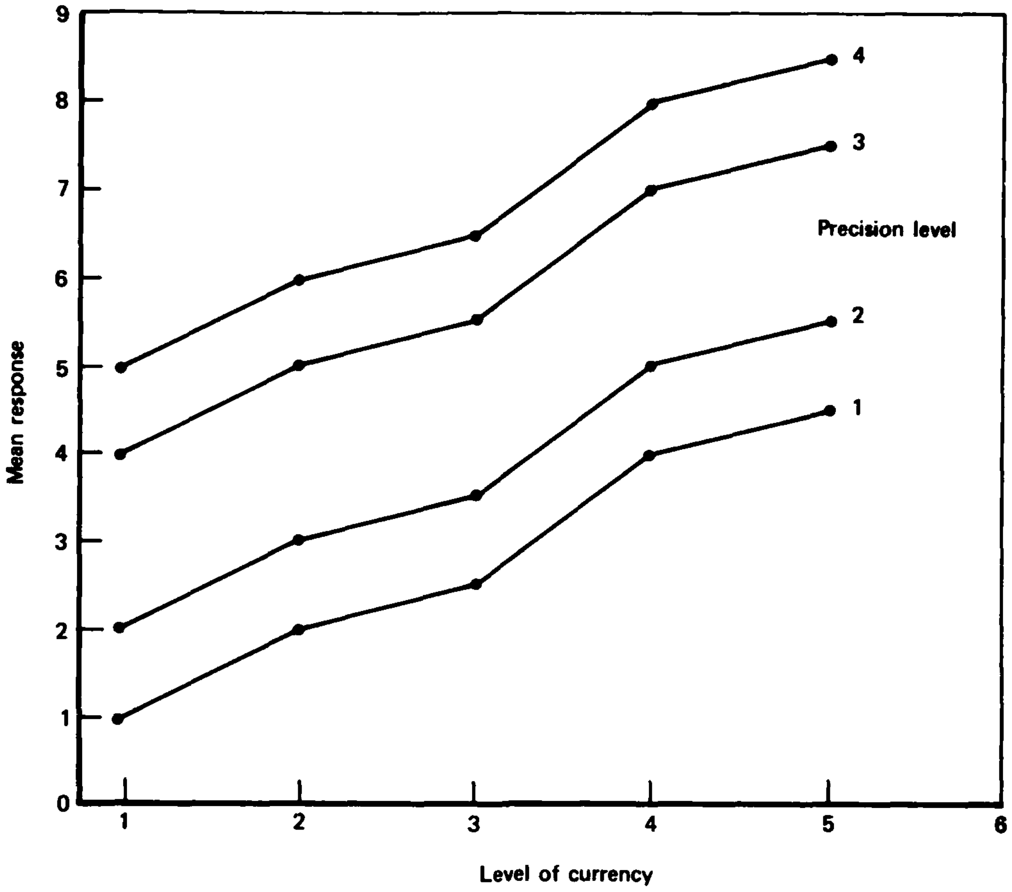

Fig. 3 — Plot of hypothetical data that support an additive model

Precision, jth level of Amount, and kth level of Currency would follow the form,

$$R_{P_iA_jC_k} = J\left[\frac{w_0 s_0 + w_P s_{P_i} + w_A s_{A_j} + w_C s_{C_k}}{w_0 + w_P + w_A + w_C} + \omega(s_{max} - s_{min})\right]; \qquad (2)$$

where J, $R_{P_iA_jC_k}$, s_{P_i}, s_{A_j}, s_{C_k}, w_0, s_0, w_P, w_A, and w_C are as described above; s_{max} and s_{min} are the scale values of the highest and lowest valued stimuli, respectively, presented in the ijkth combination; and ω is an empirical constant that represents the magnitude of the range effect (referred to as the configural-weight parameter). Note that the range model reduces to a relative-weight averaging model when s_{max} is equal to s_{min} (i.e., when the stimulus information presented for judgment is of the same subjective value) or when ω is equal to zero.

This theory predicts that the extremity of stimulus information is weighted and taken into account in the judgment process, thus producing observed interactions (nonparallelisms) among the factors when J is assumed linear. The range model predicts a particular type of interaction among the components--steady divergence ($-\omega$) or steady convergence ($+\omega$) (Birnbaum, 1974; Birnbaum and Stegner, 1979).

The range model has done well in accounting for interactions obtained in several judgment domains (Birnbaum and Stegner, 1979, 1980; Rose, 1980).

Differential-Weight Averaging Model

For three pieces of information (e.g., Precision, Amount, and Currency of Enemy Information) presented in an item, the differential-weight averaging model can be written

$$R_{P_i A_j C_k} = J\left[\frac{w_0 s_0 + w_{P_i} s_{P_i} + w_{A_j} s_{A_j} + w_{C_k} s_{C_k}}{w_0 + w_{P_i} + w_{A_j} + w_{C_k}}\right], \qquad (3)$$

where the terms are as described above. In Eq. 3, each level of each component has a separate weight; that is, weight of a piece of information depends on its scale value. These additional parameters allow more flexibility in accounting for a variety of interactions. For example, this model does well in accounting for interactions that exhibit reversals in direction.

Support for this model has been found by Anderson (1971, 1974a, 1974b), Anderson and Birnbaum (1976), and Oden and Anderson (1971).

General Comments

An important goal in research is to find the simplest explanation for observed phenomena. In judgment research, this translates into seeking a model that accounts for the data with as few parameters as possible.

Of the three equations described above, the relative-weight model of Eq. 1 requires the fewest number of estimated parameters to test its predictions. However, this model cannot account for interactions among components if they are found in the data. Both Eqs. 2 and 3 account for interactions. The range model (Eq. 2) requires an estimate of only one more parameter than the relative-weight model to test its predictions. The differential-weight averaging model, however, requires estimates of almost twice as many parameters. To accept this model as appropriate, complications in the form of the interaction to be explained would have to warrant this increase in number of estimated parameters.

Once an appropriate model is determined for a set of judgments, it is possible to understand perceptual tradeoffs among the selected components. For example, the relative-weight averaging model (Eq. 1) that explains the hypothetical data shown in Fig. 3 tells us that Enemy Information having the second level of Precision and the fifth level of Currency is equal in value to Enemy Information having the third level of Precision and the second level of Currency. Understanding these sorts of tradeoffs could be important in decisionmaking.

EXPERIMENTAL DESIGNS

For each model described above in Eqs. 1 through 3, it is necessary to estimate weight and scale value parameters to assess how well the model accounts for the judgment data. Separate estimates of weight and scale value parameters are possible only when proper experimental designs are utilized (Birnbaum, 1978; Birnbaum and Stegner, 1979; Norman, 1976). The present research employed an experimental design (described in the next section) that permits estimates of both types of parameters.

SITUATIONAL BACKGROUND INFORMATION

All three judgment experiments conducted in this research related to selecting and attacking targets in the enemy second echelon area in a Korean-like land battle. Information pertaining to the hypothetical land battle was given to each respondent before the questionnaire. This background information is presented in Callero, Naslund, and Veit (1981b).

THE RESPONDENTS

Six Air Force officers--three Captains, two Lieutenant Colonels, and one Colonel--at Langley Air Force Base, Virginia, served as respondents in Experiments II and III. The same Captains and Lieutenant Colonels participated in Experiment I; hence there were five respondents for that experiment.

II. EXPERIMENT I: PRECISION, CURRENCY, AND AMOUNT

The first experiment was designed to investigate how information about the first three components (independent variables, factors) of Enemy Information listed in Panel A, Fig. 2 affect judged value of the information for knowing the enemy's capabilities to conduct ground operations against friendly forces. The experimental design, RECIPE (Birnbaum, 1978), used to generate questions posed to respondents made it possible to test among the models shown in Eqs. 1 through 3 above.

METHOD

Respondents judged questionnaire items that characterized incoming Enemy Information for controlling attacks on fixed targets by its Precision, Currency, and/or Amount. Judgments were of the value of the incoming information for knowing the enemy's capabilities to conduct ground operations against friendly forces.

Independent Variables

Our objective was to provide a "real world" definition for each component and select component levels that spanned realistic ranges--from the best that might be achieved in the near future (highest level) to a realistic lowest level. The definitions and selected levels of each component are described below.

Precision. The Precision characteristic was a complex variable that contained information about both the preciseness of fixed target information and the preciseness of enemy vehicle information.

Fixed target preciseness related to the incoming data on the location, terrain, vulnerability, and status of fixed targets in the enemy second echelon area. Enemy vehicle preciseness related to incoming data concerning the location and discrimination of enemy vehicles in that area.

The five levels selected to define this complex variable are listed below:

P_1: Precise data on fixed targets; precise data on vehicles, can discriminate type of vehicles.

P_2: Precise data on fixed targets; precise data on vehicles, can discriminate emitting vehicles only.

P_3: Precise data on fixed targets; precise data on vehicles but no discrimination is possible.

P_4: Precise data on fixed targets; general data on vehicles and no discrimination is possible.

P_5: General data on fixed targets; general data on vehicles and no discrimination is possible.

Currency. Respondents were instructed to consider the Currency of the Enemy Information as reflecting both the Reporting Time Interval (time from detection of an enemy event to receipt by the system) and Observation Frequency (the frequency with which the area was observed). For example, respondents were instructed to consider very current data as reflecting both a rapid reporting time interval and very frequent observation of the area.

The five levels selected to define this variable are listed below:

C_1: Data are extremely current--says that the data you have on enemy events reach you instantaneously with the occurrence of the events.

C_2: Data are quite current.

C_3: Data are fairly current.

C_4: Data are slightly current.

C_5: Data are not at all current--says that the data you have on enemy events are about 9 hours old.

Amount. The Amount characteristic of Enemy Information was defined to reflect the proportion of enemy events (enemy objects or actions of military value) for which there are available data.

The four levels that defined the Amount variable are listed below:

A_1: Have data on all enemy events--says that there are data available on all enemy objects or actions of military value.

A_2: Have data on most enemy events.

A_3: Have data on some enemy events.

A_4: Have data on only a small proportion of enemy events--says that there are data available on only a small proportion of the enemy objects or actions of military value.

Stimuli and Design

The complete RECIPE design (Birnbaum, 1978) used in this experiment requires items to be generated from a fully crossed three-way factorial design, every possible two-way factorial design, and each component level presented alone. The fully crossed (5 x 5 x 4) design produced 100 items that contained one level of each of the components; every possible factorial pair produced (5x5) + (5x4) + (5x4) = 65 items that contained one level of two of the components; and each component level presented alone produced 5+5+4 = 14 items that contained a single component level, for a total of 179 experimental questionnaire items.

The following are examples of each of the three types of items used to describe Enemy Information:

1. Precise data on fixed targets; general data on vehicles. Data are fairly current. Have data on most enemy events.
2. Have data on only a small proportion of enemy events. General data on fixed targets; general data on vehicles.
3. Data are extremely current.

The order of the Precision, Currency, and Amount components was randomized within items.

Procedure and Task

The 179 items were printed in random order on the 15 pages forming the questionnaire booklet. Each respondent received a different page ordering. Each booklet also contained a set of instructions and 13 representative warm-up trials.

Respondents were instructed to consider each item to represent a different situation in which the component level(s) contained in the item was everything known about the Enemy Information coming into the system; the information had other characteristics, but these were not known. After imagining this hypothetical situation for an item, the respondent's task was to judge the overall value of the given component level(s) for knowing the enemy's capabilities to conduct ground operations against friendly forces. Value ratings were made on a nine-point scale from one (the given component level(s) of Enemy Information would be "*not at all valuable*" for knowing the enemy's capabilities to conduct ground operations against friendly forces) to nine (the given component level(s) of Enemy Information would be "*extremely valuable*" against friendly forces. Numbers between these two extremes were to be used for intermediate judgments.

Following the 179 experimental items, respondents completed an Importance Rating Task. This short three-item task required respondents to compare all pairs of the three components--Precision, Amount, and Currency--in terms of relative importance. These ratings were made on an 11-point scale. Respondents were instructed to use a one if the first component presented appeared to be very very much more important than the second, a six when the two components appeared equal in importance, and an eleven when the component presented second appeared very very much more important than the first. Again, numbers in between were to be used for intermediate judgments.

Respondents worked at their own pace, completing the task between 60 and 90 minutes. An oral and visual presentation of background information preceded distribution of the questionnaire booklets.

RESULTS AND DISCUSSION

One respondent was eliminated from the data analyses because his judgments exhibited numerous violations of the fundamental algebraic axioms of commutativity and transitivity. A series of data analyses was performed on the remaining four respondents' judgments to assess the "goodness" of the hypothesized components for inclusion in the representation and to test among the three models (Eqs. 1 through 3) hypothesized to explain the judgment data.

We wanted to group respondents for data analyses because group analyses provide a more reliable basis for conclusions. The criterion for grouping respondents was agreement on their importance ratings of the three components.

Importance Ratings

For these data, all respondents ordered the importance of the three factors the same. Currency was rated as the most important factor, Precision was rated as second in importance, and Amount was rated as least important of the three.[1]

Because of this agreement in the components' rated order of importance, all four respondents were grouped for data analyses. Analyses were also performed on individual respondent data as a check to see if the individual performed like the group.

Hypothesized Models

The models described in Eqs. 1 through 3 were each examined in terms of how well they accounted for the 179 mean judgments (averaged over all four respondents). The fit of each model was assessed assuming a linear relationship between observed and subjective responses (i.e., a linear J in Eqs. 1 through 3).

---

[1] The mean rating of the relative importance of Currency versus Precision was 4.40, a rating below the neutral value of 6, indicating that respondents considered the first factor presented for comparison, Currency, to be more important than the second, Precision, in a Currency-Precision comparison. The mean rating of the relative importance of Amount versus Currency was 8.00, a rating above the neutral value of 6, indicating that the second factor presented, Currency, was considered to be of greater importance than the first factor, Amount. The mean rating of the relative importance of Precision versus Amount was 4.20, indicating that Precision was considered to be of greater importance than Amount.

Relative-Weight Averaging Model. The relative-weight averaging model of Eq. 1 predicts parallel curves such as those shown in Fig. 3 for all combinations of components. Graphs of group and individual respondent data revealed systematic interactions between the components. These obtained interactions are inconsistent with the parallelism predictions of the relative-weight model.

Range Model. The program RECIPE (Stegner and Birnbaum, 1979) that utilizes the STEPIT subroutine (Chandler, 1969) was used to assess the fit of the range model of Eq. 2 to the data by taking the least-squares estimates of the model's parameters based on the 179 mean judgments.

A graphic test of the range model is shown in Fig. 4. In Panel A of Fig. 4, mean response is plotted as a function of the estimated scale value for the five levels of Currency, with a separate curve for each level of Precision. (Levels 2 and 3 of the Precision factor were less than 0.3 apart in scale value and were therefore combined.) In Panel B, the mean response is plotted as a function of the estimated scale value for Precision with a separate curve for each level of Amount; and in Panel C, the mean response is plotted as a function of the estimated scale value for Amount with a separate curve for each level of Currency. In each panel the curves represent the predictions of the range model; the obtained mean judgments are shown as solid points. Each data point in Panel A is the average of 20 judgments, that is, the average over all of the Precision x Currency subdesigns.[2] Each data point in Panel B is the average of 24 judgments (the average over all of the Amount x

---

[2] The P x C judgments from the three-way design were averaged over 4 levels of Amount; these were averaged with the P x C judgments from the two-way design for each of the four respondents.

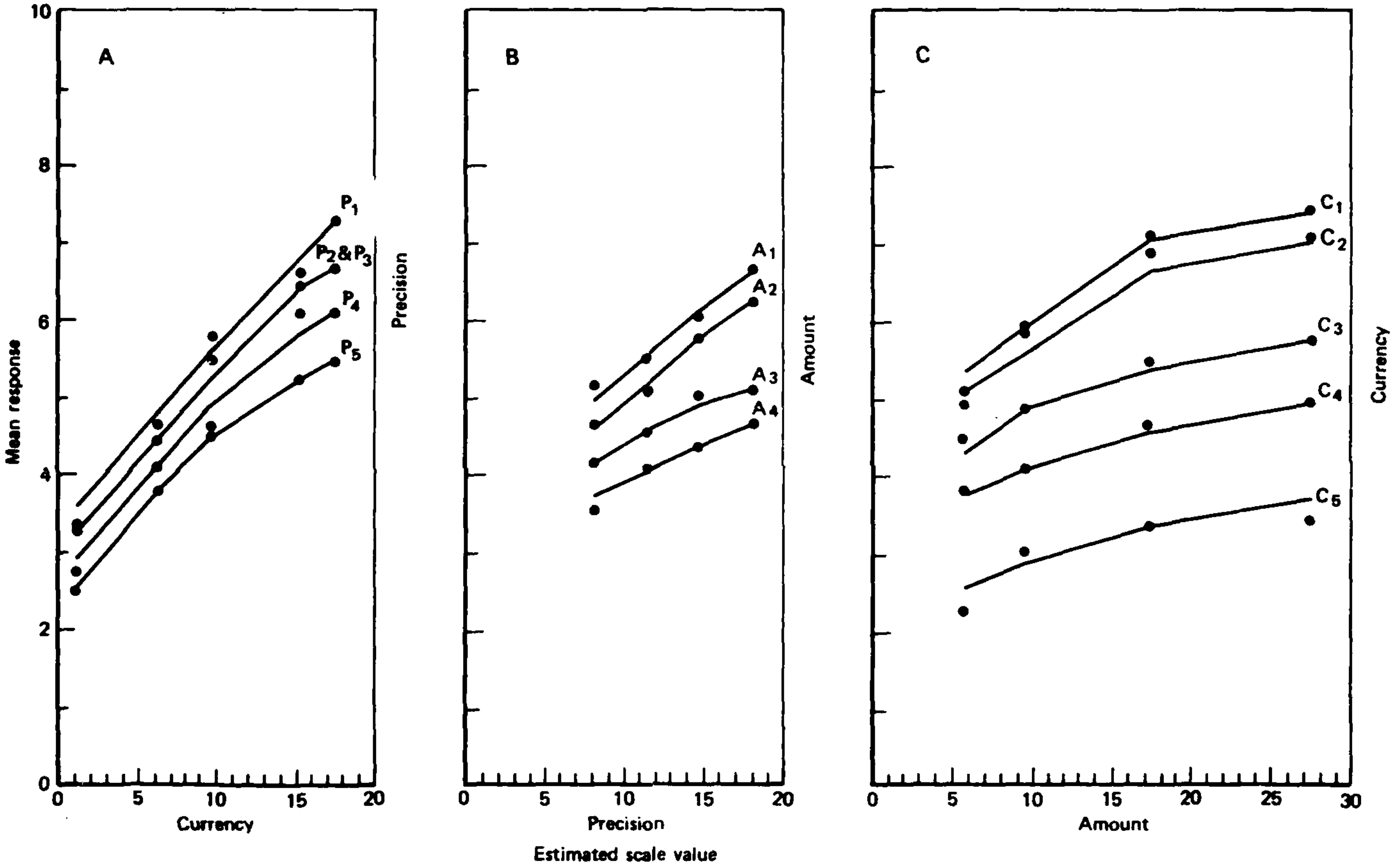

Fig. 4—Graphical assessment of the range model

In each panel the curves represent the predictions for the range model; the mean judgments are shown as solid points

Precision subdesigns).[3] Similarly, each data point in Panel C is the average of 24 judgments.

It is evident from comparing the data points with the predicted values (the curves) that the data points are very close to the model predictions. These data support the range model of Eq. 2 as an explanation of the value judgments. The least-squares solution yielded an average data-model discrepancy of 0.413 response scale points.[4] This would be considered a good fit, especially in view of the small size of the respondent sample.[5]

Since the range model did well in accounting for the data, the main effects of each of the components on judgments can be seen by examining the curves in Fig. 4. The slopes of the curves in each panel represent the effect of the abscissa component; separations between the curves depict the main effect of Precision, Amount, and Currency in Panels A, B, and C, respectively. The systematic divergent interaction can be seen in all panels by comparing the vertical separations between the curves at the highest abscissa value with the vertical separations at the lowest abscissa value. The first set of vertical separations is

---

[3] The A x P judgments from the three-way design were averaged over the five levels of Currency; these were averaged with the A x P judgments from the two-way design for each of the four respondents.

[4] The average data-model discrepancy was computed as follows. The least-squares solution yielded a sum of squared discrepancies which, when divided by the number of judgments (179), gives the average squared discrepancy; taking the square root of the average squared discrepancy yields an average data-model discrepancy.

[5] A fit of the relative-weight averaging model to the data produced a larger sum of squared errors than did the range model--38.43 for the relative-weight averaging model compared to 30.57 for the range model. Only one additional parameter needs to be estimated to test the range model, and for these data the fit is improved. From the curves in Fig. 4 (the predicted values), it appears that this improved fit results from picking up the systematic variance in the divergent interaction.

substantially greater than the second in each panel of Fig. 4. The range model accounts for this divergent interaction effect with a negative value for the configural-weight parameter (ω in Eq. 2). This divergent interaction reflects a belief by respondents that differences due to changes in one component (e.g., Precision in Panel A, Fig. 4) increase as another component (e.g., Currency in Panel A, Fig. 4) improves. A convergent interaction, produced by a positive value of the configural-weight parameter, would reflect the belief that differences due to changes in one component (e.g., Precision) decrease as another component (e.g., Currency) is improved.

As a check on how well the group data represented the individual, the range model was assessed for each respondent individually. The average data-model discrepancy was approximately 0.7 response scale points. The sign of the configural-weight parameter (ω) was negative for each of these respondents, indicating that each respondent's data followed the divergent interaction shown in Panels A to C in Fig. 4.

Differential-Weight Averaging Model. Estimates of thirteen additional parameters would be required to fit the differential-weight averaging model to these data. Because the interaction observed in the judgments was simple (i.e., a steady divergence) and the range model did a good job in accounting for the data, it was concluded that the differential-weight model would be an unnecessarily complicated explanation of these data. Thus, this model was rejected in favor of the simpler range model that accounted for the data with substantially fewer parameters.

Summary. Of the three models considered, the range model was selected as the best explanation of the simple divergent interactions observed in the data for all four respondents. Thus, this model would be considered a good candidate for the transfer function of Unit 25 in Fig. 1.

Parameters Derived from the Range Model

The fit of the range model to the data provided subjective scale values of each component level, weights of each component, and the configural-weight parameter that indicates the magnitude of the range effect on judgments.

Scale values for each level of the three components are shown in Table 1. These are the s_{P_i}, s_{A_j}, and s_{C_k} parameters in Eq. 2; the subscript i represents one of the five levels of Currency, j represents one of the five levels of Precision, and k represents one of the four levels of Amount.

Even though the Amount component of Enemy Information had only four levels, it had the largest difference in scale value between its highest and lowest level (i.e., it had the largest range in scale values among the components studied). The Precision component had the smallest range of scale values. Two Precision component levels (P_2 and P_3) received approximately the same scale value. For these respondents, being able to "discriminate emitting vehicles only" is no better than not being able to discriminate at all ("no discrimination is possible").

Table 1

SCALE VALUES DERIVED FROM THE RANGE MODEL

	Levels of Component				
Component	5	4	3	2	1
Currency	1.02	6.18	9.62	15.32	17.36
Precision	8.02	11.29	14.83	14.58	18.19
Amount		5.62	9.37	17.34	27.62

Table 2

RELATIVE WEIGHTS DERIVED FROM THE RANGE MODEL

Component	Weight
Currency	.171
Precision	.135
Amount	.097
Configural weight (ω)	-0.07

Table 2 shows the relative weights of the three components and the estimated configural-weight parameter.[6] Currency has the largest

---

[6] From Eq. 2, the relative weight of Precision is $(w_P/w_0+ w_P+ w_A+ w_C)$; the relative weight of Amount is $(w_A/w_0+ w_P+ w_A+ w_C)$; and the relative weight of Currency is $(w_C/w_0+ w_P+ w_A+ w_C)$.

relative weight, followed by Precision, with Amount having the smallest relative weight.

SUMMARY AND CONCLUSIONS

Precision, Amount, and Currency all systematically affected judgments of value for knowing the enemy's capabilities to conduct ground operations against friendly forces. This supports inclusion of these components as characteristics of Enemy Information in Unit 25 of the tactical air command and control and force employment representation shown in Fig. 1.

The fit of the range model to the data makes it a good candidate for the transfer function in experimental Unit 25 of the hierarchical representation shown in Fig. 1. This model accounted for the observed divergent interactions among these three components with a negative sign for the configural-weight parameter. The negative sign on this parameter reflects a "pessimistic" attitude toward incoming Enemy Information in the sense that low valued information is given extra emphasis (weight) in forming the overall judgment. Poor Currency, Precision, or Amount tends to pull value judgments lower than they would be if values associated with each component level were simply averaged.

III. EXPERIMENT II: PRECISION, TIMING, AND OBSERVATION FREQUENCY

Experiment II was conducted simultaneously with Experiment I as part of the effort to obtain components to define Enemy Information and best model (transfer function) to explain effects of those components on judgments. The models described in Eqs. 1 through 3 were again entertained as candidates for the transfer function. However, in this experiment, the three components hypothesized by Air Force professionals as descriptions of incoming Enemy Information for controlling attacks on fixed targets were those shown in *Panel B of Fig 2: Precision*, Timing (reporting time interval), and Observation Frequency.

In Experiment I, Currency was a single component that included both Reporting Time Interval and Observation Frequency. Experiment II investigated the behavior of "Currency" when split into its separate parts--Reporting Time Interval and Observation Frequency.

METHOD

The same experimental design used in Experiment I, RECIPE, was used to investigate how Precision, Timing, and Observation Frequency affected judgments of the value of incoming information for knowing the enemy's capabilities to conduct ground operations against friendly forces.

Independent Variables

Each of the three components shown in Panel B, Fig. 2 had five levels. Again, our objective was to define the components realistically and select component levels to span a wide range of actual possibilities.

The definitions of Precision and its five levels were the same as those used in Experiment I.

Timing or the Reporting Time Interval (RTI) was defined as the time elapsed from observation of an enemy event (enemy object or action of military value) to receipt of that information by the control function. The five levels were 0, 5 minutes, 30 minutes, 1 hour, and 3 hours.

Observation Frequency was defined as the frequency with which the enemy second echelon area was observed. The five levels of frequency were continuous, every 30 minutes, every hour, every 3 hours, and every 6 hours.

Stimuli and Design

A total of 215 questionnaire items were generated from the complete RECIPE design. The fully crossed (5 x 5 x 5) design produced 125 items, each containing one level of each of the components; every possible two-way factorial design produced 25+25+25=75 items that contained one level of two components; and 5+5+5=15 items that contained only one level of one component.

An example of each type of item is shown below:

1. Precise data on fixed targets; general data on vehicles.
   Observation Frequency is every 6 hours.
   Reporting Time Interval is 5 minutes.

2. Reporting Time Interval is 3 hours.
   General data on fixed targets; general data on vehicles.

3. Observation Frequency is continuous.

Procedure and Task

The procedure for entering the 215 items in questionnaire booklets and task instructions were identical with those described above for Experiment I. Judgments were made using the same 9-point rating scale. As in Experiment I, respondents also rated the relative importance of the three components in a three-item comparison task.

RESULTS AND DISCUSSION

Data analyses followed the same general outline as in Experiment I.

Importance Ratings

As in Experiment I, we looked for agreement in order of component importance ratings among respondents as justification for performing analyses on group data. For these importance ratings, there was little agreement on the component orders. Two respondents rated Timing as the most important, followed by Precision, with Observation Frequency rated as least important. Two respondents rated Timing as most important with Observation Frequency and Precision rated second and last, respectively. One respondent ordered Observation Frequency first, Precision second, and Timing last in importance. The importance ratings for one respondent were intransitive.[1]

---

[1] This is the same respondent that was eliminated in Experiment I. In this short importance rating task, the intransitivity was probably a result of reversing the response scale for one of the three items. Since there were only three items, however, it was not possible to ascertain where this reversal occurred.

Comment

Because of lack of agreement found in the importance ratings, model predictions were analyzed on an individual respondent basis.

Hypothesized Models

Graphs of individual respondent data showed interactions between the components for all respondents. The interactions appeared to follow a divergent pattern (as in Fig. 4) for some respondents and a convergent pattern for other respondents. Such interactions rule out a relative-weight averaging model (Eq. 1) for these data. Further, because all observed interactions exhibited simple patterns of continuous divergence or convergence, the differential-weight averaging model (Eq. 3) was rejected on the basis that it requires more parameters than needed to account for the data. The best candidate as an explanation for each respondent's data was the range model.

The parameters for the range model were estimated using all 215 judgments for each respondent separately. The average data-model discrepancy was approximately 0.9 response scale points for three of the respondents and approximately 1.1 response scale points for the other three respondents. These can be considered reasonably good fits for single-subject data with no replicates.

The order of magnitude of the scale values for the component levels was the same for all respondents. However, respondents differed in their ordering of component weights and in the sign (positive versus negative) of the configural-weight parameter, ω, that weights the range effect. Table 3 shows the component weights, the order of the component weights, and the configural-weight parameter derived from the range model for each respondent.

The six respondents produced four different orderings of model weights; respondents 1 and 2 were in agreement on the order of their model weights as were respondents 3 and 4.

The sign of the configural-weight parameter, ω, was positive for two respondents, indicating a convergent interaction among the components for these respondents, and negative for the other four respondents, indicating a divergent interaction among the components for these respondents. Thus, some of these respondents could be considered

Table 3

WEIGHTING PARAMETERS DERIVED FROM THE RANGE MODEL

	Range Model Relative Weights				
Respondent	Prec.	Timing	Obs.Freq.	Order of Model Weights	Configural Weight (ω)
1	.336	.355	.210	T > P > O	-0.133
2	.185	.285	.147	T > P > O	+0.033
3	.266	.141	.026	P > T > O	+0.078
4	.158	.085	.076	P > T > O	-0.072
5	.178	.104	.166	P > O > T	-0.120
6	.268	.238	.324	O > P > T	-0.106

P = Precision
O = Observation Frequency
T = Timing

pessimists and some of them optimists. For the pessimists ($-\omega$), if both high and low valued component levels were included in the item (e.g., Precision was good and Timing was poor), the judgment was <u>lower</u> than it would have been if the values of the component levels comprising the item had just been averaged (reflecting "Bad news, Timing is bad"). For the optimists ($+\omega$), the same combination of both high and low valued component levels produced a <u>higher</u> judgment than would be expected if the values of the component levels comprising the item had just been averaged (reflecting, "Good news, Precision is good").

IV. COMMENTS ON EXPERIMENTS I AND II

The range model shown in Eq. 2 did quite well in explaining the data for the four respondents in Experiment I. It also did well in explaining the individual-respondent data in Experiment II and would be considered the preferred model of those considered.

While respondents appear to all be represented by the same model in Experiment II, the differences in parameters shown in Table 3 present an undesirable and complicated picture. It suggests the possibility that a different set of range model weighting parameters would be needed for each respondent; that is, multiple transfer functions would be needed at Unit 25 of the command and control and force employment representation shown in Fig. 1.

The idea of multiple transfer functions within an experimental unit is unappealing and impractical. Although the intent of our ongoing research is to investigate bases for individual respondent differences, we always seek the set of components that affects judgment and results in the simplest explanation: a single transfer function for an experimental unit and the function that explains the judgment data with fewest parameters. Based on these goals, the data presented in the above two experiments suggest that the components shown in Panel A, Fig. 2 and the range model are the best component and transfer function candidates for experimental Unit 25 of Fig. 1.

V. EXPERIMENT III: THE CURRENCY OF ENEMY INFORMATION

The ultimate goal in evaluating the command and control and force employment representation shown in Fig. 1 is to understand how to change values of component levels at the lowest (input) hierarchical tier and thereby change system outcomes that those components affect. The transfer functions at this tier have to be helpful to the decisionmaker whose job it is to make these changes. Part of the practical usefulness of the transfer functions is that they specify the tradeoffs in values among the components in producing various outcomes at the experimental units. However, this will not be enough if the components are not defined along dimensions that are clearly interpretable to the decisionmaker. For example, qualitative dimensions (very much, a little) and compound definitions that include more than one dimension (e.g., Precision and Currency in Experiment I) are too difficult to interpret. In the first case, the decisionmaker cannot know for sure what equipment produces "a little" or "very much" of something. In the second case, there is no way for the decisionmaker to ferret out the separate effects of two or more dimensions included in a definition; therefore, the definition is not helpful in deciding which equipment is important. Obviously, definitions of components at this input tier require special consideration. Simple and helpful definitions would be those that consist of a single physical dimension, that is, a single dimension of physical values (e.g., time, frequency, number).

Experiment III concentrates on this issue for the Currency component (Experiment I). In Experiment I, Currency was found to be a member of a set of components that satisfactorily defined Enemy Information. However, in Experiment I, Currency was defined in terms of both Reporting Time Interval (RTI) and Observation Frequency. This compound definition does not provide information about how to change Currency (RTI and/or Observation Frequency) and thus change the judged value of Enemy Information for controlling attacks on fixed targets.

To investigate the differential effects of RTI and Observation Frequency on Currency of Enemy Information, it is necessary to manipulate RTI and Observation Frequency and have respondents judge the Currency of Enemy Information. This suggests creating a new experimental unit at the Currency position in the representation, as illustrated in Fig. 5.[1] In the new experimental unit, RTI and Observation Frequency are the independent variables hypothesized to affect the Currency of Enemy Information.

In the transfer function approach, a judgment model (transfer function) is sought to explain the relationship among the components at each experimental unit in an hierarchical representation. Thus, a transfer function needs to be sought to explain the relationship among the components that make up this new experimental unit. To get a powerful test of judgment models that might appropriately explain this relationship requires more than two components (factors) to be manipulated in the stimulus design (e.g., as in the RECIPE design used in Experiments I and II). However, the scale-free design (Birnbaum,

---

[1] The representation needs to be adjusted in this way whenever components at the existing lowest tier are defined in terms that do not allow assessment of how to change the components and thus alter outcomes they affect.

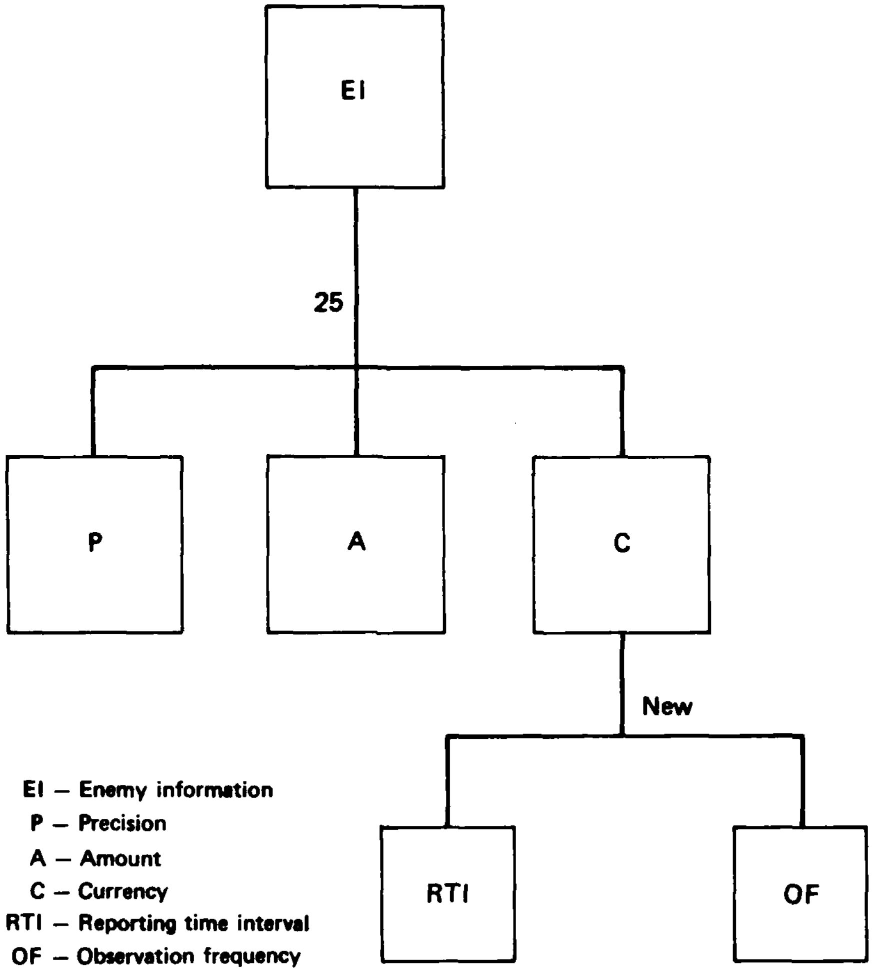

Fig. 5—Extension of Unit 25 of Fig. 1. The additional complex system hypothesis is that RTI and OF affect Currency of Enemy Information

1974; Birnbaum and Veit, 1974b) described and illustrated in this section, can be used to assess the perceptual tradeoffs among selected levels of these components.

RECIPE AND SCALE-FREE DESIGNS

The RECIPE and scale-free designs offer different contributions to judgment research. As described and illustrated in the last two experiments, the RECIPE design can provide stringent tests of hypothesized judgment models. However, these tests rely on an assumption about the form of the function relating observed to subjective responses (J in Eqs. 1, 2, and 3). In Experiments I and II, we assumed that this function was linear. Based on this assumption, we used the observed responses to test among the three models described in Eqs. 1 to 3. This led us to the conclusion that the range model did a good job in accounting for the interactions observed in the responses. However, had we assumed J in Eqs. 1 to 3 to be only monotonic, we would probably have concluded that the relative-weight averaging model of Eq. 1 was the appropriate model because a monotone transformation on the data would probably have eliminated the observed interactions (produced the predicted parallel curves). The dilemma of whether or not data should be transformed to eliminate observed interactions cannot be resolved with the RECIPE design.

The scale-free design (described below) offers a solution to this dilemma. This design provides a way to obtain the subjective responses directly. If interactions among variables are observed in the scale-free subjective responses, it would be necessary to posit a model that could account for interactions among those variables; it would be inappropriate to transform the interactions away. However, the design does not contain the constraints necessary to test judgment models that specify the nature of the effects observed in the scale-free data.

Conclusions are strengthened when both RECIPE and scale-free designs are used in the same judgment experiment, which is the plan for on going research on the tactical air command and control and force employment representation shown in Fig. 1. The logic behind the scale-free design is presented next.

THE SCALE-FREE DESIGN

As mentioned above, scale-free designs provide a way to assess perceptual tradeoffs among values placed on component levels; scale-free perceptual tradeoffs are a test of the validity of observed interactions. This is equivalent to testing the appropriateness of an additive model (e.g., the relative-weight averaging model of Eq. 1) to explain scale-free subjective responses. If the additive model (which predicts no interactions among the components) is rejected in a scale-free test of the components, it would be concluded that these components subjectively interact in affecting judgments.

A scale-free test of additivity between variables (e.g., RTI and Observation Frequency) is accomplished by embedding two combination tasks in a single task. For example, in Experiment III, each item in the questionnaire was comprised of two descriptions of Enemy Information. Each description contained a level of RTI and a level of Observation Frequency. The respondents' task was to compare the two descriptions and judge how much more current the first description was than the second. This compound task required respondents to first combine the RTI and Observation Frequency information contained in each description before comparing the two descriptions. The second comparison task allows a scale-free test of an additive model for the first task of combining values associated with RTI and Observation Frequency. This can be seen from the following logic. If comparison responses are ordinally related to the predictions of a subtractive model, then the scale values derived from the subtractive model are the subjective values associated with the Currency descriptions; these provide a scale-free test of an additive model to explain the relationship between RTI and Observation Frequency. That is, suppose that

$$R_{ijkl} = J\,[\,(o_i + r_j) - (o_k + r_l)\,], \qquad (4)$$

where R_{ijkl} is the overt response; o_i and r_j are the ith and jth levels, respectively, of Observation Frequency and Reporting Time Interval that make up the first description, and o_k and r_l are the kth and lth levels of Observation Frequency and Reporting Time Interval that make up the second description; and J is some monotone function. If comparison responses are monotonically related to the predictions of a subtractive model (i.e., if the responses fit the model directly, or a monotone transformation can be found to fit the responses to the model), scale values of the Enemy Information descriptions (each description contains a level of RTI and Observation Frequency) derived from the subtractive model provide the scale-free data; these values are the subjective responses associated with the RTI and Observation Frequency combinations under the subtractive model. A plot of these subjective responses as a function of the levels of one component (Reporting Time Interval) with a separate curve for each level of the other component (Observation Frequency) provides a test of the perceptual structure of the RTI and Observation Frequency effects on Currency judgments. If Eq. 4 is correct and RTI and Observation Frequency combine additively, the resulting curves will be parallel as in Fig. 3. This can be seen from the following development of Eq. 4. For any two levels of Observation Frequency, o_1 and o_2 and a given level of Reporting Time Interval, r_1, Eq. 4 predicts

$$R = J\,[(o_1 + r_1) - (o_2 + r_1)]$$

$$= J\,[o_1 - o_2] \quad .$$

That is, the difference between any two levels of Observation Frequency is <u>independent</u> of the level of Reporting Time Interval. Similarly, the difference between any two levels of Reporting Time Interval is predicted to be the same, independent of the level of Observation Frequency. If Observation Frequency and Reporting Time Interval are combined in a nonadditive fashion, the difference in values placed on o_1 and o_2 should <u>depend</u> on the level of Reporting Time Interval, and the curves resulting from a graphic plot of the scale-free subjective responses would reveal interactions. If an interaction between the Observation Frequency and RTI factors is concluded from the graphic plot, then removing the interaction would destroy the fit of the subtractive model. Thus, when interactions are obtained in these scale-free tests, they are considered "real" or perceptual in nature and require an explanation.

Thus, by embedding the combination process of special interest in another task (e.g., a comparison task), it is possible to distinguish between additive and nonadditive models for the combination task of interest. In Experiment III, we get a test of an additive combination rule for RTI and Observation Frequency (and hence an assessment of the perceptual tradeoffs between these two components on Currency judgments) by assuming only that comparison responses are ordinally related to subjective differences.

METHOD

In Experiment III, respondents compared the relative Currency between two descriptions of incoming Enemy Information. Thus, Currency was the dependent variable (the judged dimension).

Independent Variables

The three levels of Reporting Time Interval (R) were:

R_1 = 0,
R_2 = 30 minutes,
R_3 = 3 hours.

The three levels of Observation Frequency (O) were:

O_1 = continuous,
O_2 = every hour,
O_3 = every six hours.

Stimuli and Design

Nine different descriptions of Enemy Information were constructed from a 3 x 3, Observation Frequency x Reporting Time Interval, factorial design. Each description consisted of one level of RTI and one level of Observation Frequency. These nine descriptions formed the columns of a larger factorial design. The four sets selected to form the rows of the design were (O_3,R_3), (O_3,R_1), (O_1,R_3), and (O_2,R_1), where the levels of Observation Frequency and RTI are as defined above. The overall design can be expressed as a (4 x 9) [Row by Column] factorial design. This design generated 36 items for judgment. Each item contained two descriptions of the Currency of Enemy Information. An example of one of these items would be:

> Reporting time interval is 30 minutes.
> Observation frequency is continuous.
>
> Observation frequency is every 6 hours.
> Reporting time interval is 3 hours.

Procedure and Task

The 36 items were printed in random order in questionnaire booklets. The respondent's task was to compare the two descriptions with respect to the Currency of the Enemy Information. For each item, respondents were instructed to make their comparisons in two stages. They were instructed to first select the description that, in their opinion, resulted in the most current information, and second to select a number from the response scale that best represented how much more current that description appeared to be.

Ratings were made on a 19-point scale. Numbers between 11 and 19 were used if the first description in the item was selected as yielding more current information than the second description; selection of a particular number indicated how much more current the first description appeared to be. Numbers from one to nine were used if the second description appeared to yield more current information than the first. If both sets appeared to result in equally current information, the rating would be a 10. A set of instructions and seven representative warm-up trials preceded the 36 experimental trials.

RESULTS AND DISCUSSION

All six respondents were grouped for these analyses.

Ordinal Test of a Subtractive Model

The computer program MONANOVA (Kruskal and Carmone, 1969) was used to find the best monotone transformation for the data to fit a subtractive model of Eq. 4. The transformed data corresponded closely to this model's predictions; deviations were relatively small in magnitude and unsystematic.

Scale-Free Data Derived from the Subtractive Model

The nine column marginal means from the 4 x 9 (Row x Column) transformed data matrix are the scale-free values for the 3 x 3 (Observation Frequency x RTI) factorial design. These values represent interval scales of the subjective responses associated with the Enemy Information descriptions under the subtractive model. In Fig. 6, these subjective responses are plotted as a function of Observation Frequency with a separate curve for each level of RTI.

The form of the curves shown in Fig. 6 clearly violates the parallelism prediction of an additive model for Observation Frequency and RTI. The curves reveal a marked systematic divergence toward the highest (best) level of Observation Frequency. When the level of one component (e.g., *Observation* Frequency) was bad, the other component had less of an effect. The divergent interaction between these components is like that predicted by the range model of Eq. 2.

The perceptual tradeoffs between Reporting Time Interval and Observation Frequency can be seen by comparing the scale-free data points. For example, when RTI is 3 hours, there is very little gained in Currency value by changing Observation Frequency from every hour to continuous. However, this change in Observation Frequency makes quite a large difference when RTI is zero. In fact, the rather steep slope of the curve for RTI equal to zero (top curve in Fig. 6) shows that every improvement in Observation Frequency has a substantial effect on subjective Currency of the Enemy Information when RTI is zero. Other subjective tradeoffs can be assessed by comparing other points on the graphs.

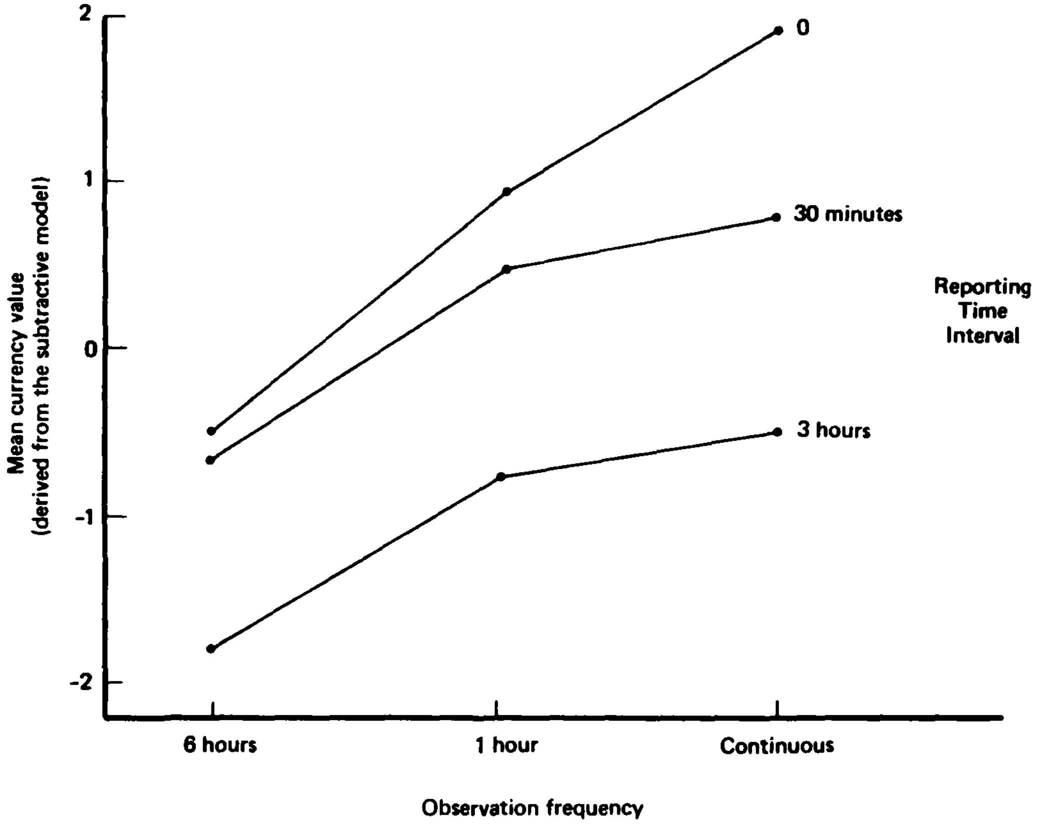

Fig. 6—Scale-free currency data

If it could be argued that the interactions observed in Fig. 6 could be rescaled to additivity, interpretation of the perceptual tradeoffs would change. However, such an argument would be inappropriate for these data; rescaling these data to additivity would require changing the rank order of the Currency comparison responses.

Comments

The scale-free design allows an examination of perceptual tradeoffs among the component levels actually manipulated in the design. As stated earlier, it does not provide the constraints to test a model that might explain the observed perceptual structure. Thus, there is no basis for inferring what the perceptual tradeoffs might be among other component levels not manipulated in the scale-free design. It is inappropriate to interpolate from the curves (see Fig. 6) because values obtained in this manner are determined in part by the spacing on the abscissa. This spacing is arbitrary since a model to account for this interaction was not tested, and thus scale values of the component levels (levels of Observation Frequency are on the abscissa) could not be determined.

VI. SUMMARY REMARKS

In this note we have presented some results of our pilot research on the tactical air command and control representation shown in Fig. 1 and demonstrated some of the measurement methodology involved in complex system analysis using the subjective transfer function approach (Veit and Callero, 1981).

In our research using the RECIPE design (Experiments I and II) to explore appropriate components and models for explaining the value of Enemy Information for controlling attacks on fixed targets (Unit 25 of Fig. 1), we found that Precision, Amount, and Currency, as defined in Experiment I, are good candidates for the components, and the range model with a negative configural-weight parameter (indicating "conservative" judgment tendencies) is a good candidate for the transfer function. Precision, Amount, and Currency were preferred as Enemy Information descriptors because judgments of these components could be explained by one model, the range model, with one set of parameters. This is the simplest of our alternative explanations. It results in but one transfer function for Unit 25 in the hierarchical representation.

In Experiment III, the scale-free design revealed a systematic divergent interaction between RTI and Observation Frequency in judgments of Currency. The scale-free design allowed assessments of the tradeoffs between these two components in their effects on perceived Currency of Enemy Information.

Conclusions concerning the appropriate components and model for Unit 25 in the hierarchical representation shown in Fig. 1 can be strengthened by incorporating Precision, Amount, and Currency in a scale-free design that provides information about the perceptual nature of the observed interactions obtained in Experiment I. Incorporating the variables used in Experiment III--RTI and Observation Frequency--in a RECIPE-like design that permits tests of models to explain the observed perceptual interaction effects between these two variables [1] is the next step needed to determine an appropriate transfer function at this new experimental unit in the hierarchy. These are considerations for future research on the hypothesized tactical air command and control representation shown in Fig. 1.

---

[1] It is interesting to note that interactions were observed between RTI and Observation Frequency in Experiment II, where judgments were about the value of incoming information for knowing the enemy's capabilities to conduct ground operations against friendly forces. For some respondents, the interaction between these two components was convergent while for others it was divergent (see Fig. 6). The scale-free data of Experiment III do not help us to interpret those interactions since the dependent variable was not the same. In order to build in more constraints for interpreting the interactions obtained in both Experiments I and II, it would be necessary to use a scale-free design that incorporated the independent and dependent variables of those experiments.

REFERENCES

Anderson, N. H., "Functional Measurement and Psychophysical Judgment," *Psychological Review*, 1970, Vol. 77, pp. 153-170.

Anderson, N. H., "Integration Theory and Attitude Change," *Psychological Review*, 1971, Vol. 78, pp. 171-206.

Anderson, N. H., "Algebraic Models in Perception," in E. C. Carterette and M. P. Freidman (eds.), *Handbook of Perception*, Vol. II, Academic Press, New York, 1974a.

Anderson, N. H., "Information Integration Theory: A Brief Survey," in D. H. Krantz, R. C. Atkinson, R. D. Luce, and P. Suppes (eds.), *Contemporary Developments in Mathematical Psychology*, Vol. 2, Academic Press, New York, 1974b.

Anderson, T. S., and M. H. Birnbaum, "Test of an Additive Model of Social Inference," *Journal of Personality and Social Psychology*, Vol. 33, 1976, pp. 655-662.

Birnbaum, M. H., "The Nonadditivity of Personality Impressions," *Journal of Experimental Psychology*, Vol. 102, 1974, pp. 543-561 (monograph).

Birnbaum, M. H., "Intuitive Numerical Prediction," *American Journal of Psychology*, Vol. 89, 1976, pp. 417-419.

Birnbaum, M. H., *A Recipe for Weights and Scale Values in Multiattribute Judgment*, University of Illinois Judgment Laboratory Technical Report No. 1, 1978.

Birnbaum, M. H., and S. E. Stegner, "Source Credibility in Social Judgment: Bias, Expertise, and the Judge's Point of View," *Journal of Personality and Social Psychology*, Vol. 37, 1979, pp. 48-74.

Birnbaum, M. H., and S. E. Stegner, "Measuring the Importance of Cues in Judgment for Individuals: Subjective Theories of IQ as a Function of Heredity and Environment," *Journal of Experimental Social Psychology*, 1980 (in press).

Birnbaum, M. H., and C. T. Veit, "Scale Convergence as a Criterion for Rescaling: Information Integration with Difference, Ratio, and Averaging Tasks," *Perception and Psychophysics*, Vol. 15, 1974a, pp. 7-15.

Birnbaum, M. H., and C. T. Veit, "Scale-Free Tests of an Additive Model for the Size-Weight Illusion," *Perception and Psychophysics*, Vol. 16, 1974b, pp. 276-282.

Birnbaum, M. H., R. Wong, and L. K. Wong, "Combining Information from Sources that Vary in Credibility," Memory & Cognition, Vol. 4, No. 3, 1976, pp. 330-336.

Callero, M., W. Naslund, and C. T. Veit, Subjective Measurement of Tactical Air Command and Control--Vol. I: Background and Approach, The Rand Corporation, N-1671/1-AF, 1981a.

Callero, M., W. Naslund, and C. T. Veit, Subjective Measurement of Tactical Air Command and Control--Vol. II: The Initial Representation, The Rand Corporation, N-1671/2-AF, 1981b.

Chandler, J. P., "STEPIT: Finds Local Minima of a Smooth Function of Several Parameters," Behavioral Science, Vol. 14, 1969, pp. 81-82.

Krantz, D. H., R. D. Luce, P. Suppes, and A Tversky, Foundations of Measurement, Academic Press, New York, 1971.

Kruskal, J. B., and F. J. Carmone, "MONANOVA: A FORTRAN-IV Program for Monotone Analysis of Variance," Behavioral Science, Vol. 14, 1969, pp. 165-166.

Norman, K. L., "A Solution for Weights and Scale Values in Functional Measurement," Psychological Review, 1976, pp. 80-84.

Oden, G. C., and N. H. Anderson, "Differential Weighting in Integration Theory," Journal of Experimental Psychology, Vol. 89, No. 1, 1971, pp. 152-161.

Rose, B. J., "The Relative Contributions of Physical, Mental, and Social Components to an Overall Measure of Health," University of California, Los Angeles, 1980 (unpublished doctoral dissertation).

Stegner, S. E., and M. H. Birnbaum, "RECIPE: A Cookbook and FORTRAN-IV Program for Weights and Scale Values," 1979 (unpublished mimeograph).

Veit, C. T., "Ratio and Subtractive Processes in Psychophysical Judgment," Journal of Experimental Psychology: General, Vol. 107, No. 1, 1978, pp. 91-107.

Veit, C. T., and M. Callero, Subjective Transfer Function Approach to Complex System Analysis, The Rand Corporation, R-2447-AF, 1981.