DOCUMENT RESUME

ED 288 912 | TM 870 736

AUTHOR	Bejar, Isaac I.
TITLE	A Psychometric Analysis of a Three-Dimensional Spatial Task.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	Office of Naval Research, Washington, D.C. Psychological Sciences Div.
REPORT NO	ETS-RR-86-19-ONR
PUB DATE	Jun 86
CONTRACT	N00014-83-C-0761
NOTE	57p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Cognitive Measurement; *Cognitive Tests; Computer Assisted Testing; *Difficulty Level; High Schools; *Latent Trait Theory; Psychometrics; *Reaction Time; Regression (Statistics); Spatial Ability; Test Items; *Visualization
IDENTIFIERS	Item Calibration; *Shepard Metzler Mental Rotations Test; *Spatial Tests

ABSTRACT

This paper considers the feasibility of incorporating research results from cognitive science into the modeling of performance on psychometric tests and the construction of test items. The paper focuses on the feasibility of modeling performance on a three-dimensional rotation task within the context of Item Response Theory (IRT). To test the feasibility of psychometrically modeling performance on this item type, an 80-item, three-dimensional rotation test was constructed using eight basic Shepard-Metzler figures. An inexpensive computer system was also developed to administer the test and record performance, including response-time data. Data were collected on high school juniors and seniors. As expected, angular disparity was a potent determinant of item difficulty. The applicability of IRT to these data was investigated by dichotomizing response time at several points and applying standard item parameter estimation procedures. It was concluded that an approach to psychometric modeling that explicitly incorporates information on the mental models that examiners use in solving an item is not only workable, but also essential for future developments in psychometrics. (Author/LPG)

A PSYCHOMETRIC ANALYSIS OF A THREE-DIMENSIONAL SPATIAL TASK

Isaac I. Bejar

This research was sponsored by the
Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research, under
Contract No. N00014-83-C-0761

Contract Authority Identification Number
NR No. 150-531

Isaac I. Bejar, Principal Investigator

Educational Testing Service
Princeton, New Jersey

June 1986

A Psychometric Analysis of a Three-Dimensional Spatial Task

Isaac I. Bejar

Educational Testing Service

Princeton, New Jersey

June 1986

This research was sponsored by the Personnel and Training
Research Programs, Psychological Sciences Division, Office of
Naval Research, under Contract No. N00014-83-C-0761, Contract
Authority Identification No. NR 150 531

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b DECLASSIFICATION / DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S) RR-86-19-ONR	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Educational Testing Service 6b. OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION Personnel and Training Research Programs Office of Naval Research (Code 1142PT) 800 North Quincy Street
6c. ADDRESS (City, State, and ZIP Code) Princeton, NJ 08541	7b ADDRESS (City, State, and ZIP Code) Arlington, VA 22217-5000
8a NAME OF FUNDING / SPONSORING ORGANIZATION 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-83-C-0761
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61153N42	RR04204	RR0420401	NR 150 531

11 TITLE (Include Security Classification)

A Psychometric Analysis of a Three-Dimensional Spatial Task

12 PERSONAL AUTHOR(S)

Isaac I. Bejar

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Technical	FROM 10/83 TO 3/85	86 June	59

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Spatial ability item response theory
05	09		cognitive psychology computerized testing

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

This paper considers the feasibility of incorporating research results from cognitive science into the modeling of performance on psychometric tests and the construction of test items. The paper focuses on the feasibility of modeling performance on a three-dimensional rotation task within the context of Item Response Theory (IRT). Three-dimensional items were chosen because there is a rich literature on the mental models that are used in their solution. To test the feasibility of psychometrically modeling performance on this item type an 80-item, three-dimensional rotation test was constructed. An inexpensive computer system was also developed to administer the test and record performance, including response-time data. Data were collected on high school juniors and seniors. As expected, angular disparity was a potent determinant of item difficulty. The applicability of IRT to these data was investigated by dichotomizing response time at several points and applying standard item parameter estimation procedures. It was concluded that an approach to psychometric modeling that explicitly incorporates information on the mental models examiners use in solving an item is not only workable, but also essential for future developments in psychometrics.

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL Dr. Charles Davis	22b TELEPHONE (Include Area Code) (202) 696-4046 22c OFFICE SYMBOL ONR 1142PT

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

Abstract

This paper considers the feasibility of incorporating research results from cognitive science into the modeling of performance on psychometric tests and the construction of test items. The paper focuses on the feasibility of modeling performance on a three-dimensional rotation task within the context of Item Response Theory (IRT). Three-dimensional items were chosen because there is a rich literature on the mental models that are used in their solution. To test the feasibility of psychometrically modeling performance on this item type an 80-item, three-dimensional rotation test was constructed. An inexpensive computer system was also developed to administer the test and record performance, including response-time data. Data were collected on high school juniors and seniors. As expected, angular disparity was a potent determinant of item difficulty. The applicability of IRT to these data was investigated by dichotomizing response time at several points and applying standard item parameter estimation procedures. It was concluded that an approach to psychometric modeling that explicitly incorporates information on the mental models examiners use in solving an item is not only workable, but also essential for future developments in psychometrics.

Acknowledgements

I'd like to express my appreciation to Min hwei Wang for performing the data analysis on this and other related reports. She also provided many useful suggestions throughout the study. Doug Jones offered valuable comments on the data; Joyce Gant developed the video version of the items and was responsible for data collection throughout much of the study; John Bittner developed the hardware device used for controlling the videodisc from the computer; Barbara Benton wrote the computer system for administration, and Dave Saxe wrote a nasty routine in assembler to allow the computer to talk to the videodisc; Bob MisLevy suggested the estimation procedure, and he and Peter Pashley offered other valuable suggestions; Lisa Jansen assisted with data collection, database management, and subject recruitment; Ka-Ling Chan was responsible for the figures; Anne Harvey performed some of the analysis; Jessie Cryer was responsible for the manuscript as well as many administrative chores related to the project. Susan Embretson and Henry Braun must be acknowledged for their valuable suggestions on an earlier draft of this study.

A Psychometric Analysis of a Three-dimensional Spatial Task

Isaac I. Bejar

Introduction

The bulk of psychometric theorizing and testing practice has been unconcerned with very detailed descriptions of the mental processes that underlie performance on test items. Instead, the focus has been on broader constructs, such as aptitude and abilities. As a result, the item, the building block for constructing a test, does not play as large a role as perhaps it should in either test construction or psychometric modeling. Thurstone, I think, deserves some of the credit, or rather some of the blame, for this situation (See Stenner, Smith, & Burdich, 1983). He was opposed to radical behaviorism and in reacting to it urged psychologists not to let the stimulus be the driving force of psychology. Although this reaction was not specifically concerned with testing, it is not hard to imagine that he might carry this perspective to the psychometric arena as well.

Whether Thurstone is responsible or not, it is accurate to say that in much of test construction the items are viewed as replicates and of little intrinsic interest. The alternative perspective is that items, far from being easily replaceable entities, are important in their own right and that accounting for differences among them with respect to their psychometric characteristics will aid our understanding of what a test measures (e.g., Bejar, 1985; Embretson, 1983), just as accounting for differences in total score variation improves our understanding of what a test measures.

Whereas the preferred methodology for modeling test score variation has been factor analysis, the methodologies of cognitive science (see

Miller, Polson, & Kintsch, 1984) appear to be suited to undertake the validation of tests from this perspective. This report focuses on mental rotation research and has the broad objective of exploring the feasibility of incorporating cognitive research with psychometric modeling and item construction. This will be investigated by examining data from a well-studied cognitive task--the three-dimensional mental rotation task--from a psychometric perspective. Specifically, the present research aims to capitalize on the significant amount of research produced in the last fifteen years in the area of spatial cognition. Much of this research has focused on the type of representation used by subjects to solve spatial problems. This progress is significant. Charles Myers, who, at ETS in the '50's, conducted much research on spatial ablity for the College Board, said that

> In this report we use the term "spatial ability" to represent a complex family of abilities with unknown interrelationships. We do not yet know of a terminology that permits a more precise and efficient language. (Myers, 1958, p. 24)

By contrast, Cooper and Shepard (1984) recently concluded after reviewing their work on mental rotation that

> In spite of some unresolved issues, the close match we have found between mental rotation and their counterparts in the physical world leads inevitably to speculations about the functions and origin of human spatial imagination. It may not be premature to propose that spatial imagination has evolved as a reflection of the physics and geometry of the external world. The rules that govern structures and motions in the physical world may, over evolutionary history, have been incorporated into human perceptual machinery, giving rise to demonstrable correspondences between mental imagery and its physical analogues. (p. 114)

In the intervening period significant research and theorizing from the factor-analytic and the experimental perspective have occurred. Much of that work has been reviewed (Corballis, 1982; Lohman, 1979; McGee, 1979). What emerges from a distilation of the literature is the presence of three mental factors; namely, spatial relations, spatial visualization, and spatial orientation. These factors have been investigated by cognitive psychologists (e.g., Pellegrino & Kail, 1982). But a specific task under the spatial relations factor has received so much attention that Corballis (1982) has raised it to the level of paradigm. That task is the three-dimensional mental rotation task. A typical stimulus used for this kind of research appears in Figure 1.

- - - - - - - - - - - - - - - -

Insert Figure 1 About Here

- - - - - - - - - - - - - - - -

The most significant finding from this line of research has been the seemingly universal finding that one feature of these stimuli, namely angular disparity, controls the response time (e.g., Cooper, 1980; Shepard & Metzler, 1971). By contrast, within psychometric settings it is usually difficult to obtain a priori predictions of the psychometric difficulty of the item, let alone of its response time. That is, within psychometrics we are often content to <u>estimate</u> difficulty. But, of course, estimating is not explaining.

It is not unreasonable to suggest that we often focus on estimating rather than explaining difficulties because of the absence of a valid psychological model for solving the item. To the extent that the psychological model is concerned with the <u>effort</u> required to solve the item it may be feasible to accurately predict an item's psychometric characteristics, especially its difficulty. If this can be done with enough precision it is in principle an alternative procedure, or at least an additional source of information, for estimating the difficulty of items. For example, it may be feasible to obtain valid estimates of difficulty by combining information about the psychological demands of the items and a small sample of subjects, instead of administering the test to a large sample of potential examinees. The implementation of this approach would require procedures for estimating the parameters in a psychometric model that are capable of incorporating "prior" information into the estimation process.

The foundation for incorporating prior information into the estimation of psychometric parameters is being laid (e.g., Bock & Atkin, 1981; Swaminathan & Gifford, 1981; Tsutakawa & Lin, 1984). Moreover, commercially available programs exist capable of handling some forms of prior information (Assessment System Corporation, 1984; Mislevy & Bock, 1982). Since the production of collateral information on the item would be based on an understanding of how examinees solve the item, the possibility also exists that at some point it might be possible to build, on the basis of that knowledge, systems whereby an item writer could receive feedback on the likely psychometric characteristics of a prospective item before it is ever administered to an examinee. (See Bejar, Stabler, & Camp 1986; Bejar & Yocom, 1986).

Psychometric Modeling of Spatial Rotation Data

The application of the foregoing to the psychometric modeling of spatial ability suggests as a criterion of success the determination of psychometric difficulty in terms of item attributes, or specifically, in the present case, linking psychometric difficulty to angular disparity. The mildest criterion, perhaps, is that difficulty should increase as angular disparity increases. A stronger criterion is that difficulty should increase in a linear fashion with angular disparity or some transformation thereof. If such a relationship can be established our interpretation of the data can be considerably more descriptive. In the absence of such a linkage, difficulty in an IRT response model is defined, when guessing is not present, as the point on the ability scale at which there is a 50-50 chance of responding correctly. When difficulty is linked to an item attribute, such as angular disparity, we can reference performance to that attribute. Thus we could speak of ability as the level required to achieve a 50-50 chance of success on a task involving a certain degree of angular disparity. In short, relating psychometric parameters to a mental model of the item solution process is likely to improve the interpretation of psychometric results.

Since the three-dimensional mental rotation item does not require problem solving, the time to obtain a correct response is directly interpretable as the efficiency with which the mental rotation takes place. Therefore, in fitting a psychometric model to these data, both accuracy and response time should be taken into consideration. Considering both responses suggests an expansion of the criterion described above. That is, in addition to expecting an increase in difficulty as a function of angular disparity, we should expect that the relationship between angular disparity and difficulty would remain the same as the time limit to perform the task is increased. Figure 2 illustrates the expected relationship. In general, however, allowance must be made for the possibility that the intercept is not linear in time. That is, in general, the gap between lines may not be a constant.

- - - - - - - - - - - - - -

Insert Figure 2 About Here

- - - - - - - - - - - - - -

Modeling Response Latency

A strategy for incorporating response latency into psychometric modeling has been proposed recently by Bloxom (1985); therefore, a review of the relevant literature will not be attempted here. Here we will focus on a discussion of modeling response latency as an extension of models for dichotomous data. The approach we follow is to fit a dichotomous item response model to response times to a set of 80 three-dimensional rotation items. The objective is to determine whether a more refined psychometric model should be attempted; not to provide the definite calibration of these data.

A common model for the probability of dichotomous response when time is not a factor is the two parameter logistic model:

$$P(u_i = 1|\theta) + \frac{1}{1 + e^{-1.7a_i(\theta - b_i)}} \qquad [1]$$

where a_i is the discrimination, b_i is difficulty parameter, and θ is ability. Now consider a situation where the interest is on the probability of a correct response after a certain period of time has elapsed. We would expect that, at least with certain item types, the longer an item is considered, the higher the probability of a correct response. Figure 3 conveys this notion. In effect we have an equation such as Equation [1] after increasing amounts of time have elapsed. Figure 3 has a constraint that is essential for interpretability, namely that the curves only differ on their inflection point. This means that the discrimination parameter is constant across time, but the difficulty parameter varies as a function of time. Put differently, the probability of a correct response after increasing elapsed times is solely a function of time. Micko (1969) has applied this idea by specifying the dichotomous item response model to be the Rasch model.

- - - - - - - - - - - - - - - -

Insert Figure 3 About Here

- - - - - - - - - - - - - - - -

Although it is in principle possible to model response time with a dichotomous item response model it is also possible to generalize the dichotomous model as Samejima (1°ʳ   as done. In this generalization the continuous response is converted to a 0-1 interval. For response latency this means that the response is expresse as a proportion of the total allowed time for responding to an item. If the time limit is 15 seconds, for example, a response latency of 5 seconds would be .33. Samejima refers to the response expressed in this manner as z. There is nothing in the model concerning whether the response time is for a correct or an incorrect response. Such a distinction must be made for scoring purposes; however, this will be discussed in a subsequent report. Basically, the idea is to treat responses as correct if, in fact, a correct response is produced after s seconds, and as not-correct-yet if an incorrect response is given. That is, incorrect responses are treated as incomplete responses indicating that "the last time we looked, namely after s seconds, the individual had not produced a correct response." In statistical terminology, incorrect responses are treated as censored observations, as in survival analysis (e.g., Miller, 1981).

With this in mind, the probability that someone of ability θ takes longer than z to respond is given by

$$P^*_{z_i}(\theta) = \int_{-\infty}^{a_g(\theta - b_{z_i})} D\exp(-Dt)\,[1+\exp(-Dt)]^{-2}dt \qquad [2]$$

$$= \left[1 + e^{-Da_i(\theta - b_{z_i})}\right]^{-1}$$

$$= \frac{1}{1 + e^{-Da_i(\theta - b_{z_i})}}$$

which is similar to Equation [1] except that the difficulty parameter now is a function of z, response time. The difficulty function b_z is not constrained to any particular shape other than that it be monotonic. In this paper we will investigate the fit of a linear function. That is, we are interested in b_z's of the form

$$b_{z_i} = \phi_i + \delta_i z \qquad [3]$$

ϕ_i could be further decomposed into components associated with figure attributes, but our focus here is on the adequacy of a linear difficulty function when angular disparity and time are taken into consideration.

An interesting implication of this model is its possible compatibility with the "slope and intercept" methodology commonly used by cognitive researchers interested in individual differences (e.g., Lansman, Donaldson, Hunt, & Yantis (1982). The slope and intercept methodology calls for computing the regression of response time on item attributes such as angular disparity, for each subject. The slope and intercept of that regression are taken to be estimates of individual differences parameters. The slope, for example, is taken to be an indicator of speed of processing while the intercept is interpreted to include a series of "overhead" processes such as encoding. It is not clear from users of this methodology what relationship ought to exist between these two parameters. However, in practice they are often correlated. In the Lansman et al. study, for example, the slope on paper and pencil three-dimensional mental rotation tests correlated .43 and .32 with the slope and intercept, respectively, of a computer version of the same test. That is, the slope and intercept seem to be picking up substantially the same individual difference variable.

The model in Equation [2] predicts that the relationship between response time and item complexity (as reflected by the difficulty parameter) for a given individual is linear and increasing for a fixed level of accuracy; that is, proportion correct. This is consistent with the "slope and intercept" methodology. However, the model also predicts that the slope of that relationship is constant across all levels of ability while the intercept varies with ability. This is not necessarily consistent with the slope and intercept methodology since the slope of the regression of response time on angular disparity has been generally interpreted as the rate at which the subjects mentally rotate the object. Thus, according to the model in Equation [2] the locus of individual difference, when accuracy is constant, is not in the rate of rotation, the slope, but rather on the intercept which is often associated with the encoding and other "overhead" processes of the item. From a substantive point of view the distinction is important since it is precisely the interpretation that subjects mentally rotate the figures that has generated so much interest in this line of research.

Design

To secure data for the study we recruited 160 high school students from a local high school. The items used in this study consisted of 80 pairs of three-dimensional Shepard-Metzler figures. The following eight figures were used (A1, A2, A3, C1, C2, C3, E1, and E2).[1] For each figure true and false pairs were constructed by rotating at angular disparities of 20, 60, 100, 140, and 180. The true pairs were constructed by rotating the same figure along the picture axis. The false pairs were constructed by rotating the mirror image instead. Altogether there were 16 items at each angular disparity. The resulting 80 items were videotaped and placed on a videodisc using the 3M mastering process.

The items were presented in two different orders. In one, the examinees saw items at 100, 60, 180, 140, and 20 degrees. At each angular disparity there were 16 possible items, and one of those was chosen at random. With the second ordering, subjects worked items of 20, 140, 180, 60, and 100 degrees. Approximately one-half of the subjects took the item in each order.

The instrumentation for each data collection station consisted of the following components:

* A 64K microcomputer with 1-disc drive (Radio Shack 26-3127)
* A videodisc player (Pioneer PR 8210)
* An Amdek Color I Monitor
* A Joystick (Radio Shack, 26-3012)
* A computer-to-videodisc interface (especially constructed for the project)

[1] The author is indebted to Professor Roger Shepard of Stanford University for providing these materials.

The microcomputer was programmed to control the videodisc as well as record the responses. Response time was recorded in "ticks" where a tick is a 60th of a second. Subjects responded by means of a joystick connected to the computer. A "yes" response was signaled by moving the joystick forward, while a "no" response was indicated by moving the joystick backward.

Because of the potential unfamiliarity of the equipment, at least as a psychological testing device, careful attention was given to the instructions. Instructions were tested with several students unrelated to the study to insure that they were fully understandable. Students were told that they were to respond by moving the joystick and that they were to respond as quickly as possible without sacrificing accuracy. The instructions appear in Appendix B.

The examinee's first task was to respond to a simple task, namely to indicate whether an arrow was pointing up or down. This was done to familiarize each examinee with the response device as well as to time their reaction time to a task with almost no cognitive load. From these responses it is possible to obtain an estimate of the motor speed of individuals. These data were not analyzed as part of this study, however.

After the arrow task, subjects were given instructions on the rotation test. As part of the instructions they were able to manipulate an animation sequence containing a true and a false item. This allowed the examinee to become familiar with true and false items at all possible angles. Also, seeing the rotation in real time, examinees may have been encouraged to use a rotation strategy to solve the items. The examinees were allowed to do six practice items. The practice items were followed by 80 real items. There was a 15-second time limit for each item. After 15 seconds a "timeout" message was given if the subject had not responded.

However, students could pace themselves in the sense that they controlled when the next item was administered. At the end of each item students were told whether they had responded correctly or incorrectly.

Parameter Estimation

With the growing interest in the psychometric modeling of response time (Bloxom, 1985; Scheiblechner, 1985; Thissen, 1983), it is likely that estimation procedures tailored to response time will be forthcoming. In the meantime it is possible to obtain estimates of item parameters through estimation procedures designed for the dichotomous case. That is the approach taken here. In a nutshell, the approach calls for successively dichotomizing response time and fitting a one-parameter logistic model at each dichotomization point. Imposing a one-parameter model across angular disparities implements the constraint that discrimination be constant across time.

Each basic item, i.e., pair of distinct figures, was fitted separately. Figure 4 shows the structure on the data matrix for a given basic item. The notation "0/1," which occurs only on the upper left corner, indicates that the entries in that section of the matrix could be 1, a correct response was given, or 0, a correct response was not given.

- - - - - - - - - - - - - -

Insert Figure 4 About Here

- - - - - - - - - - - - - -

The notation "0/1/2" indicates that responses in that block could be 0, a correct response was not given; 1, a correct response was given, or 2, the "item" was not presented. A response within this type of block was coded 2 if in a preceding block a correct response was not given to that item. For example, if at the end of the three-second interval an examinee responds "false" to a true item, then that item is coded 2 in subsequent intervals. The notation 2 simply means that all the items in that block were treated as not presented. For each basic figure there were 50 "items" corresponding to the true angular disparities and true/false classification, and 5 time intervals.

Each of the eight data matrices were analyzed separately with BILOG (Mislevy & Bock, 1982) specifying a one-parameter logistic model. The resulting estimates were rescaled with respect to the distribution of ability estimates estimated with the EAP algorithm.

Results

Unlike the typical mental rotation experiment in which subjects receive a great deal of practice time, the subjects in this study spent altogether no more than forty minutes, including instruction and practice items, on the mental rotation task. Therefore, it is important to verify that the usual finding concerning the linearity of response time on angular disparity is replicated in this case. Figure 5 shows the relationship between angular disparity and response time for correct responses for true and false items across the eight basic figures. Figure 5 suggests that there is, for the most part, a good linear fit to the data. The largest residual is at 100 degrees. Apart from this, there is relatively little scatter around the best-fitted line. In fact, the fit appears better than has been reported in some studies. The impact of angular disparity on response time is less potent for false items, as Figure 5 shows.

Insert Figure 5 About Here

To assess the fit of Equation [2] we will examine two criteria. One is the relationship of difficulty to angular disparity; the other is the relationship of reaction time and angular disparity for subjects of different ability levels. In the first instance, on the basis of Equation [2] we expect the relationship to be linear with angular disparity and to maintain the same slope as time goes by. Secondly, if Equation 2 is a valid model for these data we expect to find that for single subjects, or groups of subjects, the slope of the relationship of response time on angular disparity is constant across these subjects or groups.

Figure 6 shows the results from the item calibration for each basic item. Figure 7 shows the equivalent plots for the false items. The first expectation is largely fulfilled for the true items. That is, there is a nearly perfect linear relationship between response time and angular disparity. Moreover, the slope of the best-fitting line does not change with time. The major deviation from expectation occurs at 100 degrees. This angular disparity proved to be consistently more difficult than expected. In addition, for some items, the relationship between difficulty and angular disparity appears to be nonlinear beyond 5 seconds. For the false items, however, it is clear that angular disparity is not a determinant of response time in these data (but may be so with highly practiced subjects).

- - - - - - - - - - - - - - - - - -

Insert Figures 6 and 7 About Here

- - - - - - - - - - - - - - - - - -

Figures 8 and 9 show the mean difficulty estimates across the eight basic figures for true and false items. In Figure 8 we notice again the discrepancy at 100 degrees and a slight deviation from linearity beyond 5 seconds.

- - - - - - - - - - - - - - - - - -

Insert Figures 8 and 9 About Here

- - - - - - - - - - - - - - - - - -

The second prediction we wish to test is that the individual differences are reflected on the intercept and not the slope. To invalidate this prediction it is sufficient to show that it is not so for individual examinees or groups of examinees that differ in spatial ability. Since we did not have available an independent measure of spatial ability the approach taken here is to compare two groups widely different on SAT-M. Two groups were formed based on the upper- and lower-third scores on SAT-M. The higher group consisted of 38 examinees with a mean SAT-M of 713; the lower group consisted of 31 students with a mean of 444. Figure 10 shows the relationship between proportion correct and mean reaction time (i.e., regardless of whether the response was correct or incorrect)[2] with angular disparity for these two groups The higher SAT group has a higher slope, suggesting that they are not rotating as fast, but also a smaller intercept. Thus, other things being equal, the higher SAT examinees are faster on the items with smaller angular disparity and about as fast on the higher angular disparities. It is also true, as can be seen in Figure 10, that the accuracy rate is different, being higher for the high SAT-M group on the more discrepant items.

[2]Since the model does not specifically distinguish between correct and incorrect responses and focuses on the modeling of response time, it is more appropriate to plot response time rather than response time of correct responses in the context of model fitting. Nevertheless, the plot with mean correct response time was also done, but similar results were obtained.

Insert Figure 10 About Here

Discussion

The motivation behind this work has been both substantive and methodological. The substantive interest has been the assessment of the convergence or lack thereof of two approaches to individual differences: a psychometric approach and an approach inspired by cognitive psychology. Hunt and MacLeod (1978) have expressed concerns that the differences between these two approaches may be irreconcilable. They give as an example the tendency for psychometricians to focus on global scores, such as proportion correct, versus the tendency of the cognitive psychologist to focus on more reductionist parameters such as slope and intercepts of performance on task attributes. This paper shows that the dichotomy need not exist, at least not when we adopt IRT as a psychometric framework. For example, person characteristic curve methodology (Carroll, Meade, & Johnson, in preparation; Trabin & Weiss, 1983) is well suited to characterize subject performance in a psychologically meaningful fashion. This paper demonstrates that the IRT framework can be expanded when the response of interest is the response time and in doing so demonstrates the possibility of encompassing both global and reductionist views of individual differences within the same measurement framework.

According to Equation [2] the locus of individual differences is on the intercept and not on the slope. The results presented in Figure 10 show that the relationship of reaction time to angular disparity has a different slope <u>and</u> intercept for groups of presumably different spatial ability. If we believe that the intercept is the correct indicator of ability to rotate, then we would conclude that the higher SAT group is higher on that ability. That conclusion would be consistent with existing literature (e.g., Fennema & Sherman, 1977). If we believe that the slope is the correct indicator, we would conclude that the high SAT group is of lower ability, a finding which would not be consistent with the literature. A reconciliation of these opposing conclusions lies in an accounting of the differences between the high and low SAT-M group on their accuracy. That is, the slope of the relationship between response time and angular disparity cannot be compared unless accuracy is constant in the two groups, (cf. Kail, 1985).

According to Equation [2], the slope of the relationship between the magnitude of the response, $\underline{z}$, and difficulty can be controlled through the accuracy parameter. (See Appendix A). For a given ability level a smaller slope can be obtained by reducing accuracy as difficulty increases. A larger slope will be obtained, for example, by holding accuracy constant as difficulty increases. Therefore, the differences in slope seen in Figure 10 do not necessarily violate the prediction of the model and could well be eliminated after adjusting for accuracy. Moreover, the adjustment would not have to be done explicitly in practice since, in the estimation of ability, accuracy and speed are taken into account. That is, IRT may provide a solution to the speed-accuracy tradeoff problem (see Thissen, 1983). The estimation of ability will be treated in a subsequent report.

The second motivation for this work has been purely psychometric, and the work focuses on the development of a generative approach to psychometric modeling and test administration (Bejar & Yocom, 1986). By a generative approach I mean a methodology where the generation of the items is controlled by an algorithm encoding sufficient knowledge about the mental processes underlying performance on the item that it is capable of anticipating the psychometric characteristics of the item before it is administered to examinees. The goals of this methodology are a natural extension of adaptive testing (Weiss, 1983). In an adaptive test a computer retrieves from an existing database (containing previously calibrated items) the item that is most informative for a given individual. However, with a generative approach, an item, instead of being retrieved, is created specifically for an individual in such a way that the anticipated psychometric characteristics of the items are maximally informative for each individual.

One of the goals of adaptive testing has been the achievement of high measurement precision throughout the ability range. This is a special concern with dichotomous items where a balance must be struck between overall level of precision and distribution of precision at different levels of ability. As we move from dichotomous to continuous responses that balance takes care of itself in the sense that for continuous response models information may be high throughout the ability range and in some cases equally high at all ability levels (Samejima, 1973). Therefore, that goal of adaptive testing seems to be automatically satisfied through the use of continuous responses. Nevertheless, there may be advantages to adapting the angular disparity to individuals of different abilities as a means of insuring a common response strategy by all examinees.

The two essential ingredients in implementing a generative approach are (a) a psychometric framework of sufficient flexibility; and (b) a knowledge base about performance on the item. The three-dimensional cube item was chosen for this study because of the potent effect of angular disparity on performance in mental rotation items. It is therefore possible to establish the relationship between angular disparity and difficulty based on a few points. By manipulating this feature of the item it becomes possible to generate items of any arbitrary difficulty. In a practical implementation of this idea we may have, say, 30 basic items that can be presented at rotations ranging from 20 to 180. Although everyone would be presented all 30 items, the rotation at which they are actually presented would be different for different individuals.

The measurement model that was explored in this paper to explain performance on a three-dimensional rotation item is an extension of existing item response models in current use for the dichotomous response case (Samejima, 1983). Two predictions of this model were explored as a means of testing its fit. One prediction concerned the linearity between difficulty and angular disparity. The results suggested that the prediction was substantially satisfied. The exception was at 100 degrees which was found to be a more difficult angular disparity than expected. Also, beyond 5 seconds a nonlinear relationship appears to emerge.

Implicit in this prediction is the assumption that both true and false versions of an item would have the same slope. This was distinctly not the case. That is, it appears that a two-dimensional model would be required to provide an adequate description of the true and false data. (To assess the bias that may have been introduced into the item parameter estimates as a result of the bi-dimensionality, the model was fitted to true items only, but no noticeable difference could be seen in the resulting estimates.)

The second prediction was concerned with the interpretation of slope and intercept parameters. According to the model, individuals of different ability have the same slope but different intercepts. The results suggested that considering the different accuracy rates of a high- and low-scoring SAT-M group, the difference in slopes could not necessarily be interpreted as valid indicators of rotation speed. In effect, our results suggest that the bulk of individual differences is on the intercept rather than the slope parameter. This does not necessarily conflict with research suggesting that rotation speed is an important individual-differences variable since examinees were exposed to he items for a relatively short period of time. Indeed, in contradiction of other research, the males and females studied do not differ on the slopes and intercepts (Bejar & Harvey, in progress). Whereas sex differences in tasks involving mental rotation appear t. .e a well established fact (Linn & Petersen, 1985), and because of the relatively brief exposure that subjects had, our data may not be typical.

While results of the study show that the idea of generating items of arbitrary difficulty is indeed feasible, there are some difficulties that must be borne in mind with even relatively simple stimuli, such as the three-dimensional rotation items. Specifically, performance on false items is not a function of angular disparity; on the surface this finding therefore suggests that performance on the false items is controlled by a different combination of mental processes (see Carter, Pazak, & Kail, 1983). In short, further psychometric examination of the three-dimensional cubes could focus on multidimensional modeling of processing and decision processes and on models that characterize the change in performance as a function of practice.

References

Assessment Systems Corporation. (1984). Users Manual for the MicroCat Testing System. Minneapolis, MN.

Bejar, I. I. (1985). Speculations on the future of test design. In S. E. Embretson (Ed.), Test design: Contribution from psychology, education and psychometrics (pp 279-294). NY: Academic Press.

Bejar, I. I., & Harvey, A. (in progress). Mental rotation and performance on the SAT-M. Princeton, NJ: Educational Testing Service.

Bejar, I. I., Stabler, E., & Camp, R. (1986). Syntactic complexity and psychometric difficulty: A preliminary investigation (Research Report). Princeton, NJ: Educational Testing Service.

Bejar, I. I., & Yocom, P. (1986). Analysis and generation of hidden figure items: A cognitive approach to psychometric modeling (RR-86-19). Princeton, NJ: Educational Testing Service.

Bloxom, B. (1985). Consideration in psychometric modeling of response time. Psychometrika, 50, 383-397.

Bock, R. D., & Atkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of the EM algorithm. Psychometrika, 46, 443-459.

Carroll, J. B., Meade, A., & Johnson, E. S. (in preparation). Test analysis with the person characteristic function: Implications for defining abilities.

Carter, P., Pazak, B., & Kail, R. (1983). Algorithms for processing spatial information. Journal of Experimental Child Psychology, 36, 284-304.

Cooper, L. A. (1980). Spatial information processing: Strategies for research. In R. E. Snow, P. A. Federico, & W. E. Montague (Eds.), Aptitude, learning and instruction. Vol. 1: Cognitive processes analyses of aptitudes (pp 149-176). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cooper, L. A., & Shepard, R. N. (1984, December). Scientific American, 251(6), 106-114.

Corballis, M. C. (1982). Mental rotation: Analysis of a paradigm. In M. Potegal (Ed.), Spatial abilities: Developmental and psychological foundations. NY: Academic Press.

Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. Psychological Bulletin, 93, 175-197.

Fennema, E., & Sherman, J. (1977). Sex related differences in mathematics achievement, spatial visualization and affective factors. American Educational Research Journal, 14, 51-71.

Hunt, E., & MacLeod, C. M. (1978). The sentence verification paradigm: A case study of two conflicting approaches to individual differences. Intelligence, 2, 129-144.

Kail, R. (1985). Development of mental rotation: A speed accuracy study. Journal of Experimental Child Psychology, 40, 181-192.

Lansman, M., Donaldson, G., Hunt, E., & Yantis, S. (1982). Ability factors and cognitive processes. Intelligence, 6, 347-386.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. Child Development, 56, 1479-1498.

Lohman, D. F. (1979). Spatial ability: A review and reanalysis of the correlational literature. Technical Report No. 8, Aptitude Research Project, Stanford University.

McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal and neurological influences. Psychological Bulletin, 86, 889-918.

Micko, H. C. (1969). A psychological scale for reaction time measurement. Sctu Psychologica, 30, 324-335.

Miller, J. R., Polson, P. G., & Kintsch, W. (1984). Problems of methodology in cognitive science. In W. Kintsch, J. R. Miller, P. G. Polson (Eds.), Methods and tactics in cognitive science. Hillsdale, NJ: Lawrence Erlbaum.

Miller, R. G. (1981). Survival analysis. New York: Wiley.

Mislevy, R. J., & Bock, R. D. (1982). BILOG: Item analysis and test scoring with binary logistic models. Mooresville, IN: Scientific Software.

Myers, C. (1958). Some observations of problem solving in spatial relations tasks. ETS Research Bulletin RB58-16. Princeton, NJ: Educational Testing Service.

Pellegrino, J. W., & Kail, R. (1982). Process analysis of spatial aptitude. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (Vol. 1). Hillsdale, NJ: Lawrence Erlbaum.

Samejima, F. (1973). Homogeneous case of the continuous response model. Psychometrika, 38, 203-219.

Samejima, F. (1983). A general model for the homogeneous case of the continuous response. (Report No. ONR/RR-83-3). Knoxville, TN; University of Tennessee.

Scheiblechner, H. (1985). Psychometric models for speed-test construction: The linear experiential model. In S. E. Embretson (Ed.), Test design: Developments in psychology and psychometrics. New York: Academic Press.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. Science, 171, 701-703.

Stenner, A. J., Smith, M., & Burdich, D. S. (1983). Toward a theory of construct definition. Journal of Educational Measurement, 20, 305-316.

Swaminathan, H., & Gifford, J. A. (1981). Bayesian estimation in the three-parameter logistic model. Research Report LR-119. School of Education, University of Massachusetts.

Thissen, D. (1983). Timed testing: An approach using item response theory. In D. J. Weiss (Ed.), New horizon in testing. New York: Academic Press.

Trabin, T. E., & Weiss, D. J. (1983). The person response curve: Fit of individuals to item response theory models. In D. J. Weiss (Ed.), New horizons in testing: Latent trait test theory and computerized adaptive testing (pp 83-108). New York: Academic Press.

Tsutakawa, R. K., & Lin, H. Y. (1984). Bayesian analysis of item response curves. Research Report 84-1. Department of Statistics, University of Missouri, Columbia, MO.

Weiss, D. J. (Ed.). (1983). New horizons in testing. New York: Academic Press.

Figure 1

Sample True and False Three-Dimensional Rotation Items

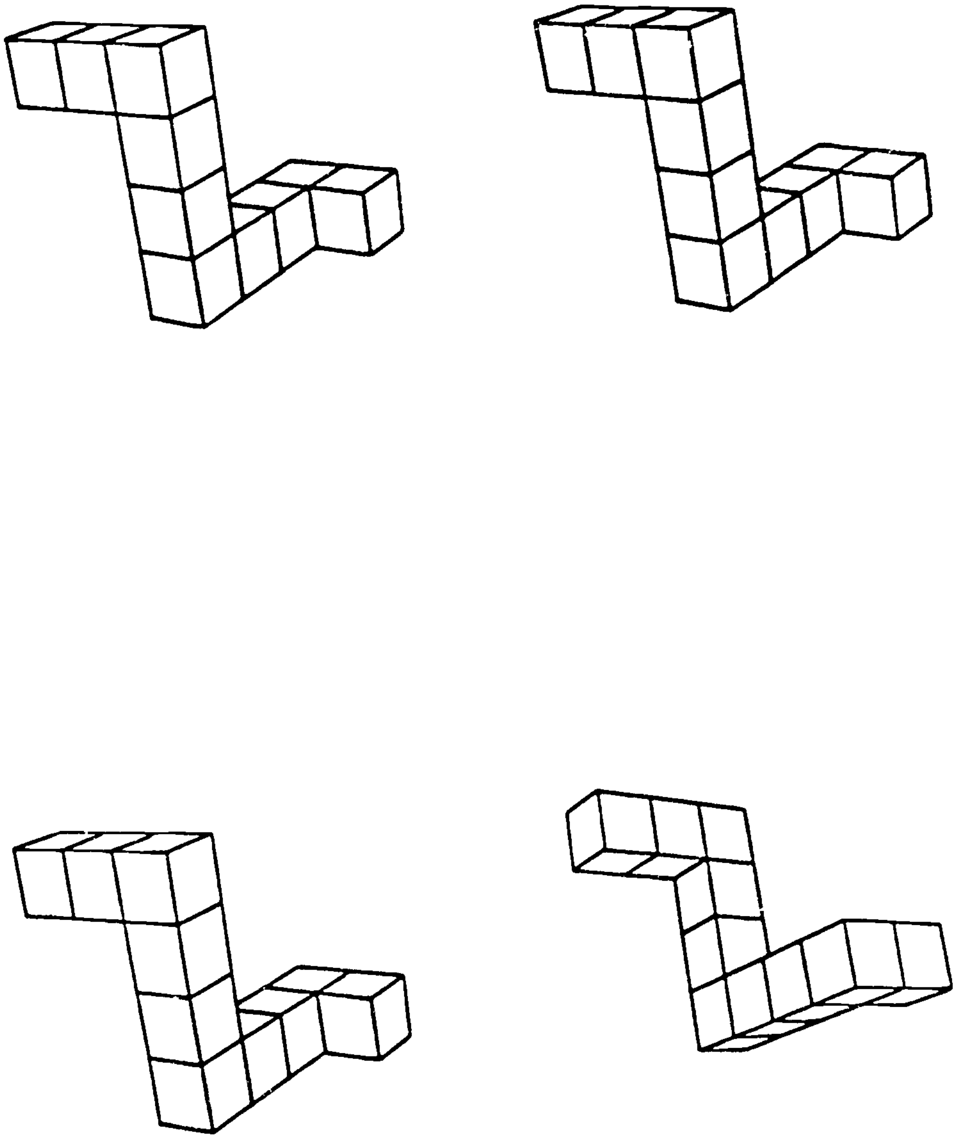


Figure 2

Hypothetical Relationship Between Difficulty

and Angular Disparity and Function of Time

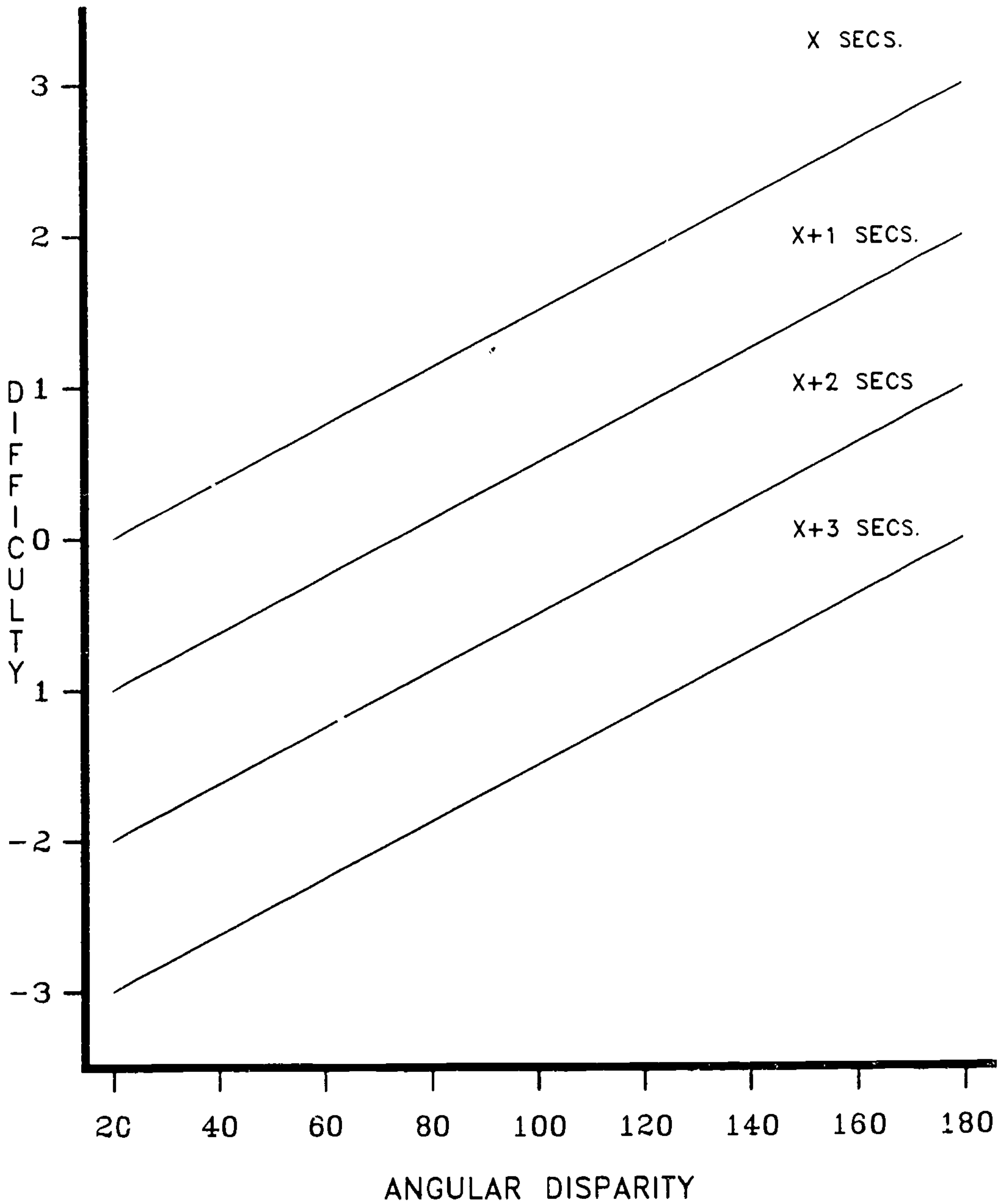


Figure 3

Hypothetical Item Response Functions as a Function of Time

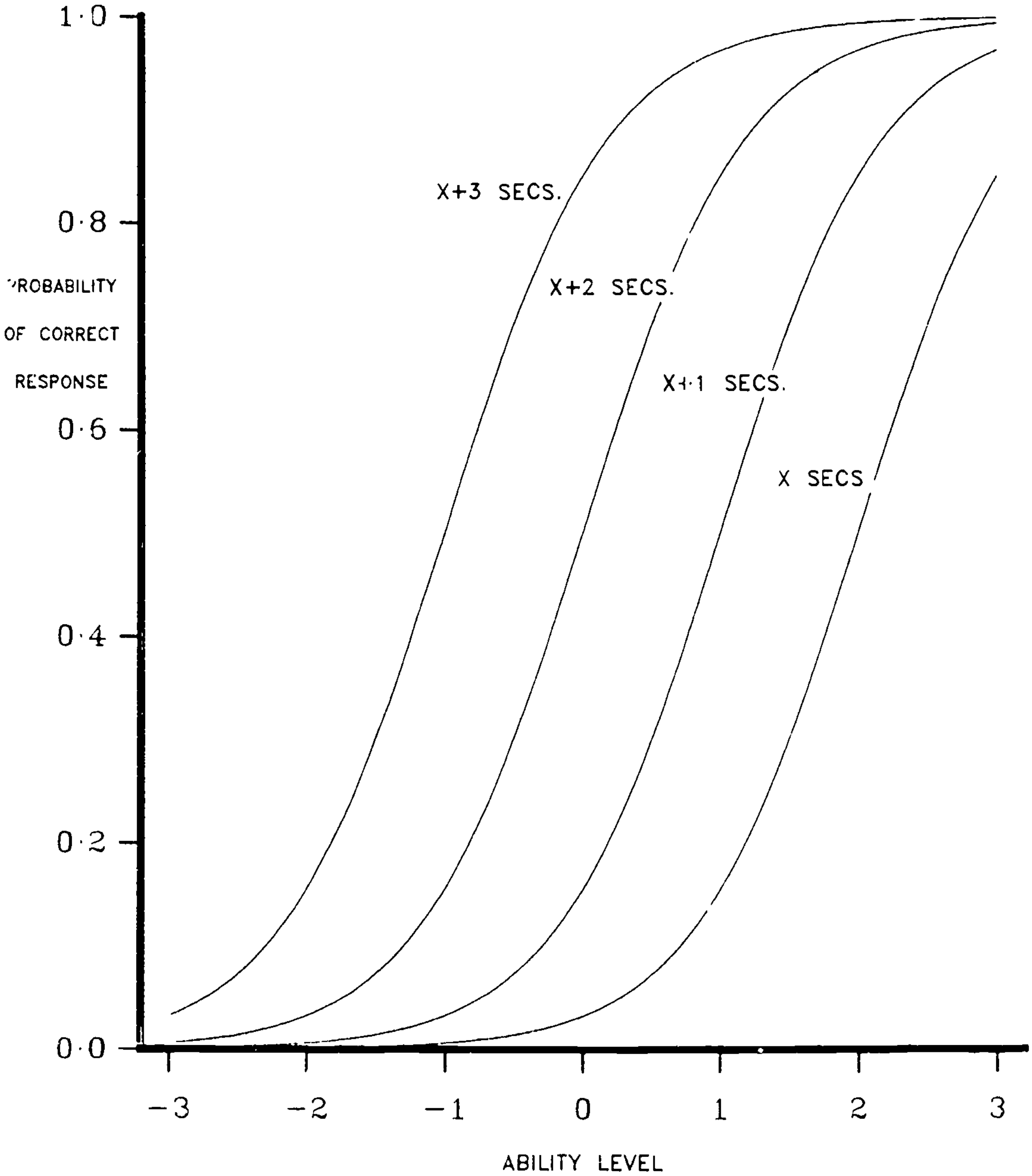


Figure 4

Data Matrix for True and False Versions of an Item at Successive Dichotomization Points and Angular Disparities

	Seconds	3 Seconds 20° T F · 60° T F · 100° T F · 140° T F · 160° T F	4 Seconds 20° T F · 60° T F · 100° T F · 140° T F · 180° T F	5 Seconds 20° T F · 60° T F · 100° T F · 140° T F · 180° T F
Subjects 1	3	0/1	2	2
	4	2	0/1/2	2
	5	2	2	0/1/2
	6	2	2	2
	7	2	2	2
Subjects 2	3	0/1	2	2
	4	2	0/1/2	2
	5	2	0/1/2	0/1/2
	6	2	2	2
	7	2	2	2

Figure 5

Relationship Between Reaction Time and Angular Disparity for True and False Versions Averaged Across the Eight Items

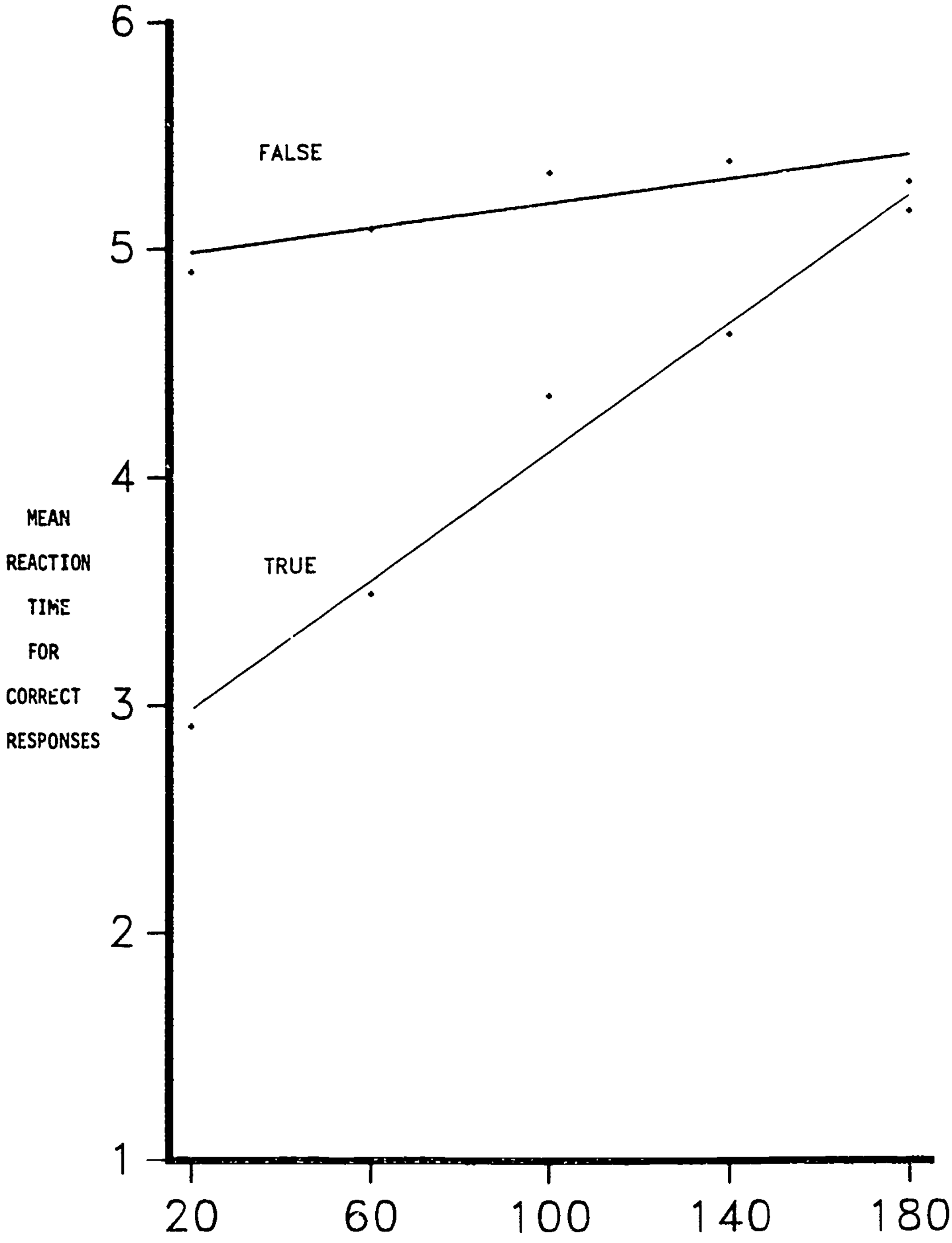
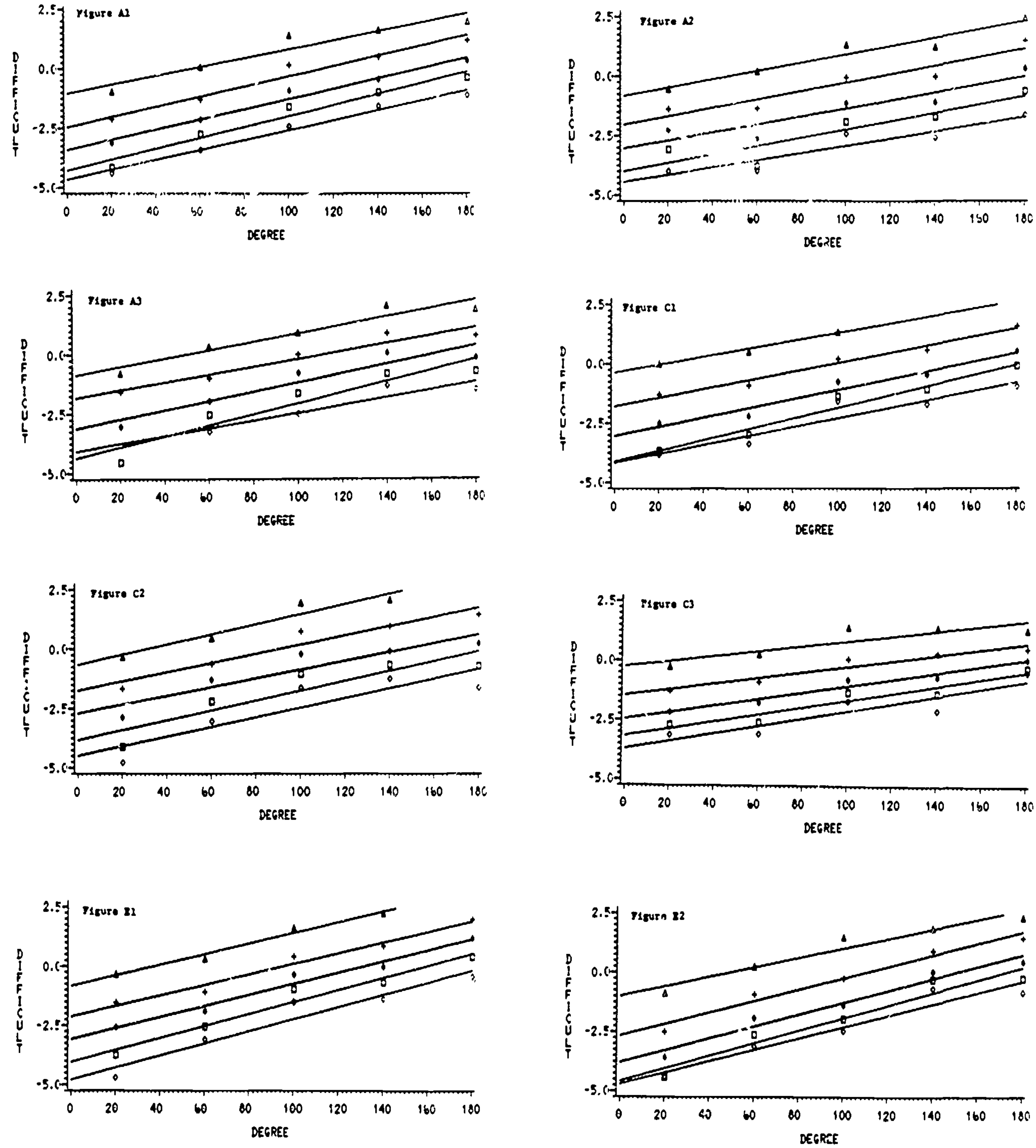


Figure 6

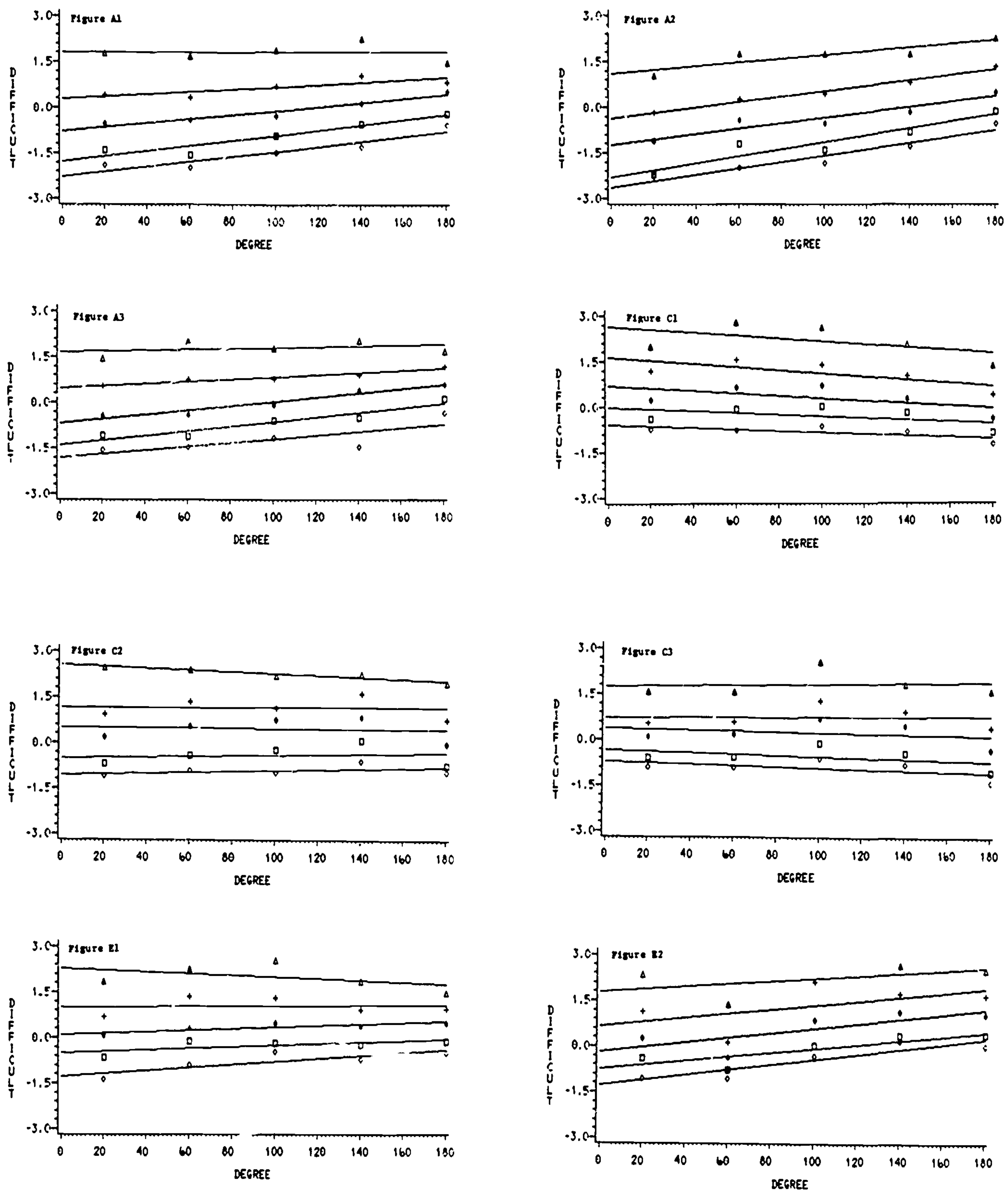
Relationship Between Difficulty and Angular Disparity for Each of the Eight Basic Items as a Function of Time (True Version)



TRIANGLE= 3 SECONDS
PLUS = 4 SECONDS
STAR = 5 SECONDS
SQUARE = 6 SECONDS
DIAMOND= 7 SECONDS

Figure 7

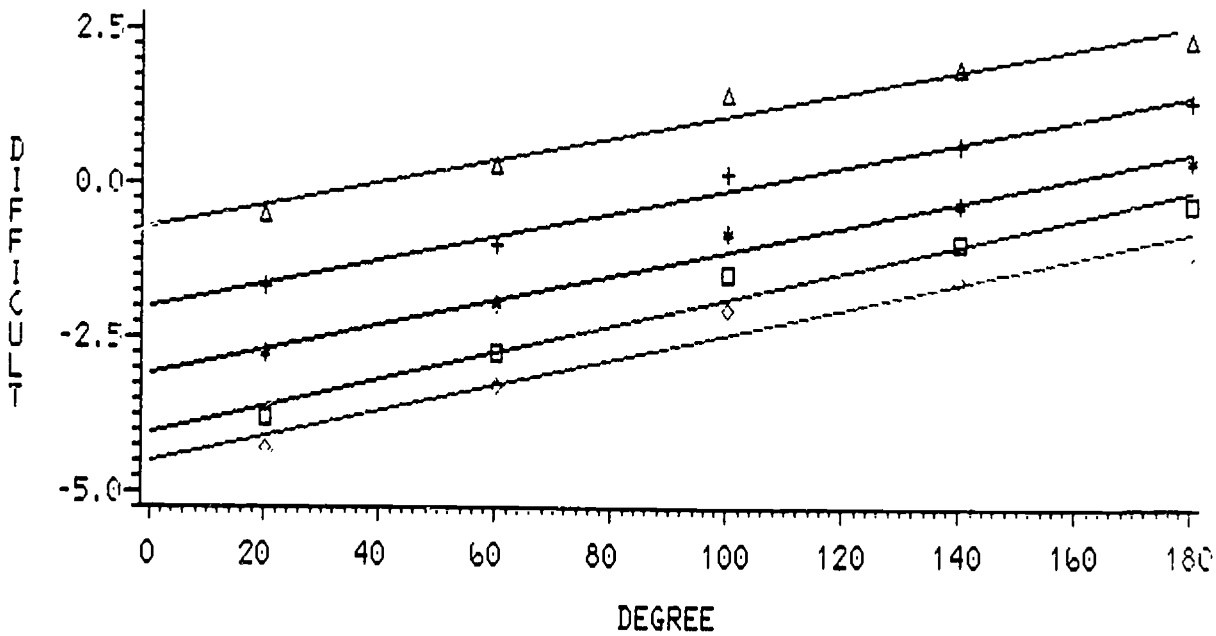
Relationship Between Difficulty and Angular Disparity for Each of the Eight Basic Items as a Function of Time (False Version)



TRIANGLE= 3 SECONDS
PLUS = 4 SECONDS
STAR = 5 SECONDS
SQUARE = 6 SECONDS
DIAMOND= 7 SECONDS

Figure 8

Relationship Between Difficulty and Angular Disparity as a Function of Time Averaged Across All Items (True Version)

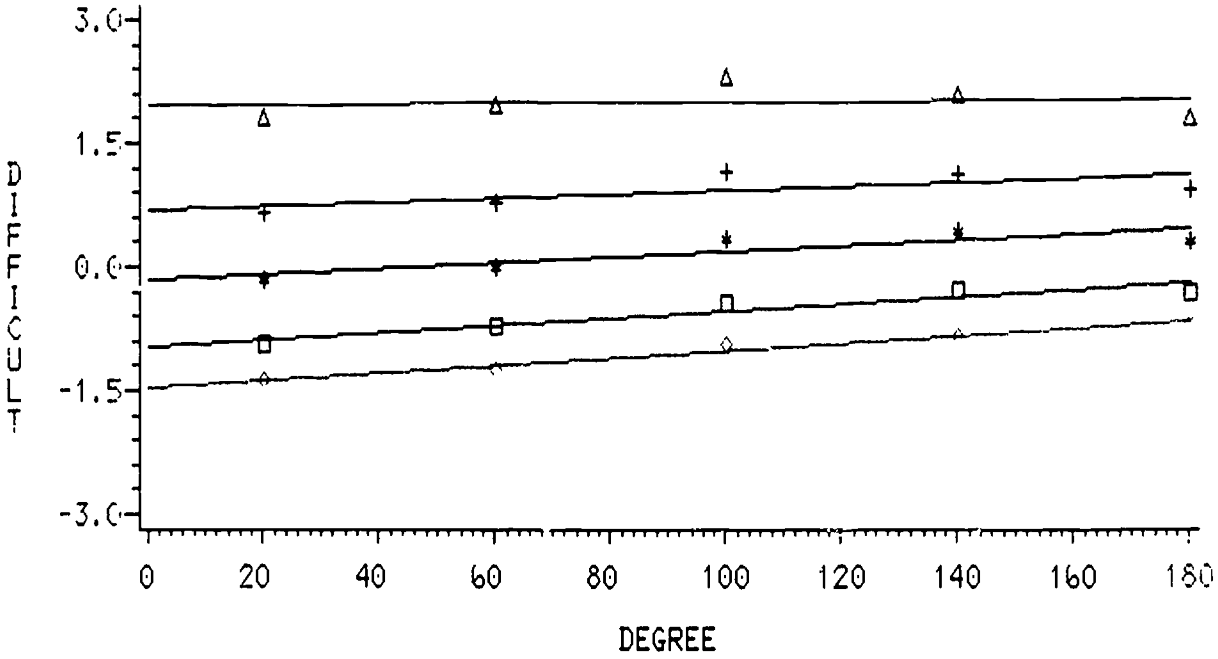


TRIANGLE= 3 SECONDS
PLUS = 4 SECONDS
STAR = 5 SECONDS
SQUARE = 6 SECONDS
DIAMOND= 7 SECONDS

Figure 9

Relationship Between Difficulty and Angular Disparity as a

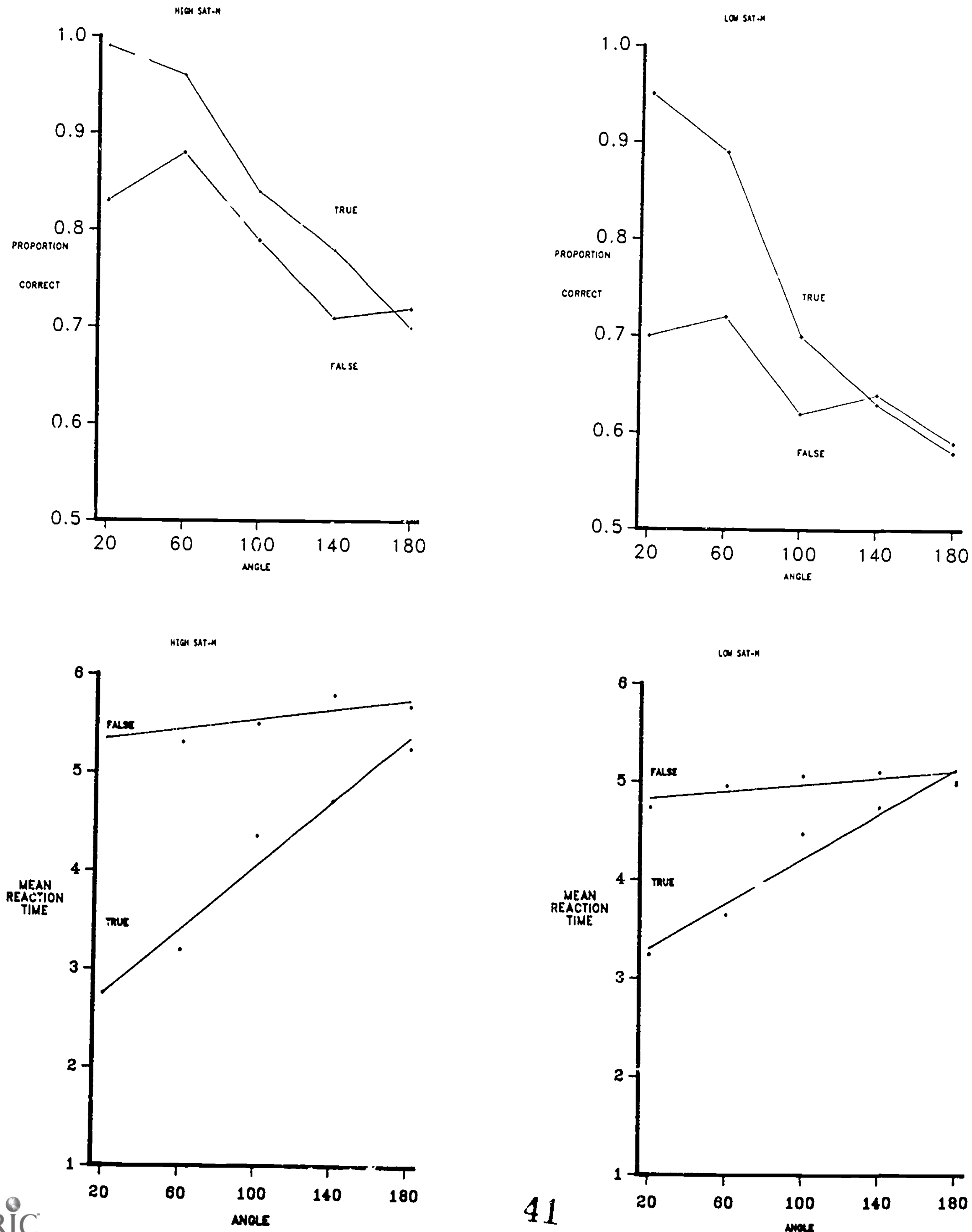
Function of Time Averaged Across Eight Items (False Version)



TRIANGLE= 3 SECONDS
PLUS = 4 SECONDS
STAR = 5 SECONDS
SQUARE = 6 SECONDS
DIAMOND= 7 SECONDS

Figure 10

Relationship Between Proportion Correct and Reaction Time as a Function of Angular Disparity for High and Low Scoring SAT-M Groups



APPENDIX A

To explore the implications of the model it is a matter of substituting values for the parameters. In this appendix we will show the relationship between reaction time and difficulty and show how the slope of the relationship between the response and difficulty can be manipulated by changing accuracy.

The response model is

$$P_z^* = \frac{1}{1 + e^{-Da_i(\theta - b_{z_i})}} \qquad [A1]$$

We will interpret P_z^* as the accuracy and z as the response time. We wish to observe the response time as a function of accuracy. (Actually, since the model is oriented such that a large response is associated with higher θ we will study the relationship with $i-z$ instead.) For illustration purposes we will assume there are four items of the same discrimination but different difficulty and will examine the results of the model for ability level $\theta = 1$. More concretely, assume:

$a_i = 1.00$ for $k=\underline{1}..4$

$b_{z_k} = 3z + \phi_k$ $(k=\underline{1}..4)$

$\phi_k = .25k - .25$ $(k=\underline{1}..4)$

$\theta = 1.0$

To simulate different response modes we will vary P_z^* and observe what happens to $(1-z)$. The results for three "response modes" appear in Table A1. Column A has the results for a careless mode. That is the response $(1-z)$ remains constant as difficulty increases but accuracy decreases from .60 to .30. For mode C, however, (1-z) increases from .75 to .99, and accuracy remains constant. That is, in order to maintain a c· .nstant accuracy rate on increasingly difficult items, 1-z must increase.

Table A1

	A		B		C	
k	(1-z)	P^*	(1-z)	P^*	(1-z)	P^*
1	.75	.60	.75	.60	.75	.60
2	.75	.50	.79	.55	.83	.60
3	.75	.40	.83	.50	.91	.60
4	.75	.30	.88	.45	.99	.60

Figure A1 plots those data. As can be seen the slope is zero for A and highest for C. In short, the slope of reaction time on item attributes is partly a [illegible]ion of the accuracy with which the individual chooses to respond.

- - - - - - - - - - - - - - -

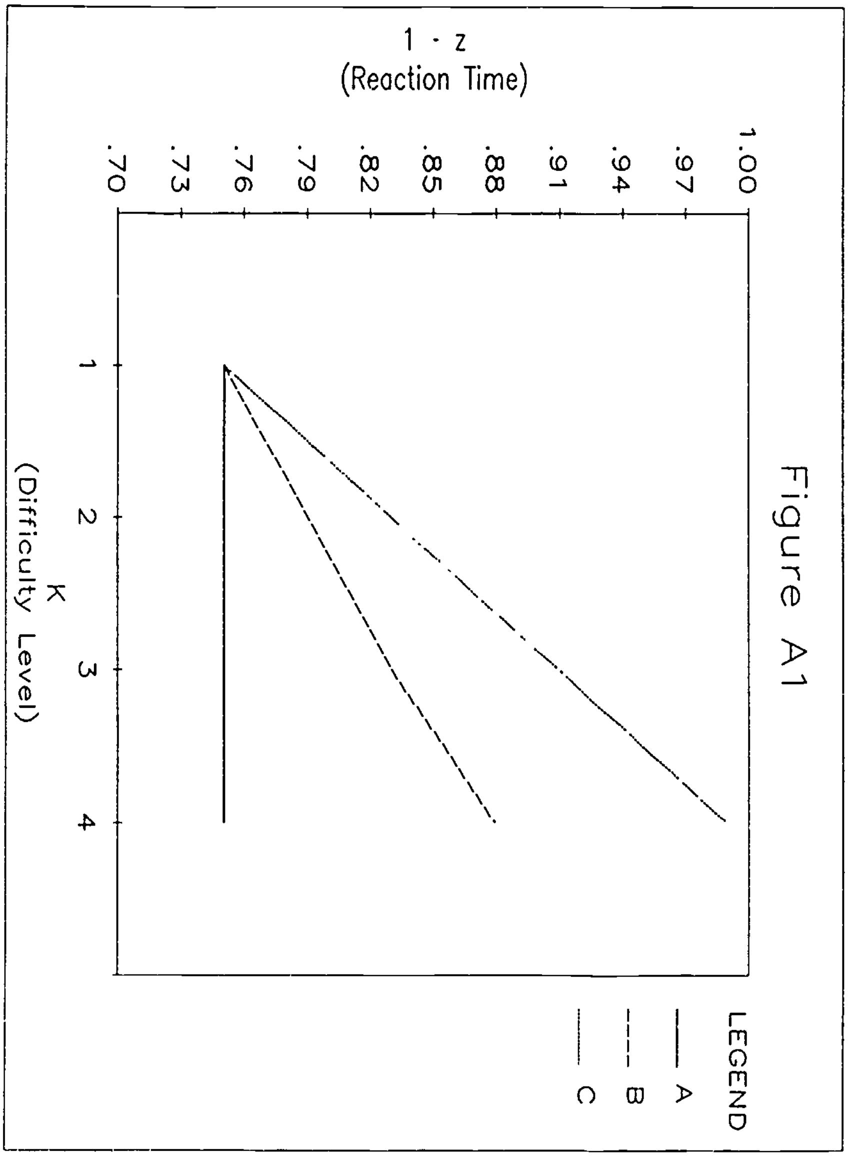
Insert Figure A1 About Here

- - - - - - - - - - - - - - -

Figure A1

Relationship Between Response Time and Difficulty

Level for Three Levels of Accuracy



APPENDIX B

Instruction for Mental Rotation Items

In the following exercises you must respond with the joystick and the red button. To get you used to responding in this way we want you to practice on some samples. In these samples you will see an arrow pointing up or down:

PUSH the stick FORWARD if arrow points UP.

PULL the stick BACK if arrow points down.

RESPOND AS QUICKLY AS YOU CAN.

(Forty arrow items are presented.)

The first kind of exercise you will work on consists of two figures. Your task is to decide whether or not the two figures are the same. It is important to be FAST and CORRECT. Sometimes the two figures will be the same even though the one on the right may be drawn at a different angle. Sometimes the two figures will not be the same.

To show you what we mean, in the following example the two figures are the same but are at a different angle. To see the example press the red button. Each time you press the red button the figure on the right will move closer to the one on the left.

(The subject is able to rotate the figure back and forth.)

If you would like to see this example again push the joystick forward. To see the rest of the instructions press the red button.

You just saw an example of figures that are the same. In the following example the two figures are NOT the same. To see the example press the red button. Each time you press the red button the figure on the right will come closer to the one on the left but they will still not be the same.

If you would like to see this example again push the joystick forward. To see the rest of the instructions press the red button.

First you will take some practice trials. It is important to be FAST and CORRECT. However, you can pace yourself because with the red button you control when to see the next trial. The time you take between trials is not counted.

Respond QUICKLY and CORRECTLY.

PUSH joystick FORWARD

if figures are the SAME.

PULL joystick BACKWARD

if the figures are NOT THE SAME.

Press the red button when you are ready for the next trial.

(Six practice items are presented.)

You are now ready for the real trials. Remember!

Respond QUICKLY and CORRECTLY.

PUSH joystick FORWARD

if figures are the SAME.

PULL joystick BACKWARD

if the figures are NOT THE SAME.

Press the red button when you are ready for the next trial.

(The 80 items are next)

Personnel Analysis Division,
AF/MPXA
5C360, The Pentagon
Washington, DC 20330

Air Force Human Resources Lab
AFHRL/MPD
Brooks AFB, TX 78235

Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

Dr. Erling B. Andersen
Department of Statistics
Studiestraede 6
1455 Copenhagen
DENMARK

Dr. Phipps Arabie
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Technical Director, ARI
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Eva L. Baker
UCLA Center for the Study
of Evaluation
145 Moore Hall
University of California
Los Angeles, CA 90024

Dr. Isaac Bejar
Educational Testing Service
Princeton, NJ 08450

Dr. Menucha Birenbaum
School of Education
Tel Aviv University
Tel Aviv, Ramat Aviv 69978
ISRAEL

Dr. Arthur S. Blaiwes
Code N711
Naval Training Systems Center
Orlando, FL 32813

Dr. Bruce Bloxom
Administrative Sciences
Code 54B1
Navy Postgraduate School
Monterey, CA 93943-5100

Dr. R. Darrell Bock
University of Chicago
NORC
6030 South Ellis
Chicago, IL 60637

Cdt. Arnold Bohrer
Sectie Psychologisch Onderzoek
Rekruterings-En Selectiecentrum
Kwartier Koningen Astrid
Bruijnstraat
1120 Brussels, BELGIUM

Dr. Robert Breaux
Code N-095R
Naval Training Systems Center
Orlando, FL 32813

Dr. Robert Brennan
American College Testing
Programs
P. O. ox 168
Iowa City, IA 52243

Dr. Patricia A. Butler
OERI
555 New Jersey Ave., NW
Washington, DC 20208

Mr. James W. Carey
Commandant (G-PTE)
U.S. Coast Guard
2100 Second Street, S.W.
Washington, DC 20593

Dr. James Carlson
American College Testing
Program
P.O. Box 168
Iowa City, IA 52243

Dr. John B. Carroll
409 Elliott Rd.
Chapel Hill, NC 27514

Dr. Robert Carroll
OP 01B7
Washington, DC 20370

Dr. Norman Cliff
Department of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

Director,
Manpower Support and
Readiness Program
Center for Naval Analysis
2000 North Beauregard Street
Alexandria, VA 22311

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Dr. Hans Crombag
University of Leyden
Education Research Center
Boerhaavelaan 2
2334 EN Leyden
The NETHERLANDS

CTB/McGraw-Hill Library
2500 Garden Road
Monterey, CA 93940

Dr. Dattprasad Divgi
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. Hei-Ki Dong
Ball Foundation
800 Roosevelt Road
Building C, Suite 206
Glen Ellyn, IL 60137

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC
(12 Copies)

Dr. Stephen Dunbar
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. James A. Earles
Air Force Human Resources Lab
Brooks AFB, TX 78235

Dr. Kent Eaton
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. John M. Eddins
University of Illinois
252 Engineering Research
Laboratory
103 South Mathews Street
Urbana, IL 61801

Dr. Susan Embretson
University of Kansas
Psychology Department
426 Fraser
Lawrence, KS 66045

ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

Dr. Benjamin A. Fairbank
Performance Metrics, Inc.
5825 Callaghan
Suite 225
San Antonio, TX 78228

Dr. Leonard Feldt
Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Dr. Richard L. Ferguson
American College Testing
Program
P.O. Box 168
Iowa City, IA 52240

Dr. Gerhard Fischer
Liebiggasse 5/3
A 1010 Vienna
AUSTRIA

Prof. Donald Fitzgerald
University of New England
Department of Psychology
Armidale, New South Wales 2351
AUSTRALIA

Mr. Paul Foley
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Carl H. Frederiksen
McGill University
3700 McTavish Street
Montreal, Quebec H3A 1Y2
CANADA

Dr. Robert D. Gibbons
University of Illinois-Chicago
P.O. Box 6998
Chicago, IL 69680

Dr. Janice Gifford
University of Massachusetts
School of Education
Amherst, MA 01003

Dr. Robert Glaser
Learning Research
& Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Bert Green
Johns Hopkins University
Department of Psychology
Charles & 34th Street
Baltimore, MD 21218

Dr. Ronald K. Hambleton
Prof. of Education & Psychology
University of Massachusetts
at Amherst
Hills House
Amherst, MA 01003

Ms. Rebecca Hetter
Navy Personnel R&D Center
Code 62
San Diego, CA 92152-6800

Dr. Paul W. Holland
Educational Testing Service
Rosedale Road
Princeton, NJ 08541

Prof. Lutz F. Hornke
Institut fur Psychologie
RWTH Aachen
Jaegerstrasse 17/19
D-5100 Aachen
WEST GERMANY

Dr. Paul Horst
677 G Street, #184
Chula Vista, CA 90010

Mr. Dick Hoshaw
OP-135
Arlington Annex
Room 2834
Washington, DC 20350

Dr. Lloyd Humphreys
University of Illinois
Department of Psychology
603 East Daniel Street
Champaign, IL 61820

Dr. Steven Hunka
Department of Education
University of Alberta
Edmonton, Alberta
CANADA

Dr. Huynh Huynh
College of Education
Univ. of South Carolina
Columbia, SC 29208

Dr. Robert Jannarone
Department of Psychology
University of South Carolina
Columbia, SC 29208

Dr. Douglas A. Jones
P.O. Box 6640
Lawrenceville
NJ 08648

Dr. G. Gage Kingsbury
Portland Public Schools
Research and Evaluation Department
501 North Dixon Street
P. O. Box 3107
Portland, OR 97209-3107

Dr. William Koch
University of Texas-Austin
Measurement and Evaluation
Center
Austin, TX 78703

Dr. Leonard Kroeker
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Michael Levine
Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat 23
9712GC Groningen
The NETHERLANDS

Dr. Robert Linn
College of Education
University of Illinois
Urbana, IL 61801

Dr. Robert Lockman
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. Frederic M. Lord
Educational Testing Service
Princeton, NJ 08541

Dr. James Lumsden
Department of Psychology
University of Western Australia
Nedlands W.A. 6009
AUSTRALIA

Dr. William L. Maloy
Chief of Naval Education
and Training
Naval Air Station
Pensacola, FL 32508

Dr. Gary Marco
Stop 31-E
Educational Testing Service
Princeton, NJ 08451

Dr. Clessen Martin
Army Research Institute
5001 Eisenhower Blvd.
Alexandria, VA 22333

Dr. James McBride
Psychological Corporation
c/o Harcourt, Brace,
Javanovich Inc.
1250 West 6th Street
San Diego, CA 92101

Dr. Clarence McCormick
HQ, MEPCOM
MEPCT-P
2500 Green Bay Road
North Chicago, IL 60064

Mr. Robert McKinley
University of Toledo
Department of Educational Psychology
Toledo, OH 43606

Dr. Barbara Means
Human Resources
Research Organization
1100 South Washington
Alexandria, VA 22314

Dr. Robert Mislevy
Educational Testing Service
Princeton, NJ 08541

Headquarters, Marine Corps
Code MPI-20
Washington, DC 20380

Dr. W. Alan Nicewander
Department of Psychology
University of Oklahoma
Oklahoma City, OK 73069

Dr. William E. Nordbrock
FMC-ADCO Box 25
APO, NY 09710

Dr. Melvin R. Novick
356 Lindquist Center
for Measurement
University of Iowa
Iowa City, IA 52242

Director, Manpower and Personnel
Laboratory,
NPRDC (Code 06)
San Diego, CA 92152-6800

Library, NPRDC
Code P201L
San Diego, CA 92152-6800

Commanding Officer,
Naval Research Laboratory
Code 2627
Washington, DC 20390

Dr. James Olson
WICAT, Inc.
1875 SouthState Street
Orem, UT 84057

Office of Naval Research,
Code 1142PT
800 N. Quincy Street
Arlington, VA 22217-5000
(6 Copies)

Special Assistant for Marine
Corps Matters,
ONR Code 00MC
800 N. Quincy St.
Arlington, VA 22217-5000

Dr. Judith Orasanu
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Wayne M. Patience
American Council on Education
GED Testing Service, Suite 20
One Dupont Circle, NW
Washington, DC 20036

Dr. James Paulson
Department of Psychology
Portland State University
P.O. Box 751
Portland, OR 97207

Dr. Roger Pennell
Air Force Human Resources
Laboratory
Lowry AFB, CO 80230

Dr. Mark D. Reckase
ACT
P. O. Box 168
Iowa City, IA 52243

Dr. Malcolm Ree
AFHRL/MP
Brooks AFB, TX 78235

Dr. Carl Ross
CNET-PDCD
Building 90
Great Lakes NTC, IL 60088

Dr. J. Ryan
Department of Education
University of South Carolina
Columbia, SC 29208

Dr. Fumiko Samejima
Department of Psychology
University of Tennessee
Knoxville, TN 37916

Mr. Drew Sands
NPRDC Code 62
San Diego, CA 92152-6800

Dr. Robert Sasmor
HQDA DAMA-ARL
Pentagon, Room 3E516
Washington, DC 20310-0631
USA

Dr. Mary Schratz
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. W. Steve Sellman
OASD(MRA&L)
2B269 The Pentagon
Washington, DC 20301

Dr. Kazuo Shigemasu
7-9-24 Kugenuma-Kaiga
Fujusawa 251
JAPAN

Dr. William Sims
Center for Naval Analysis
4401 Ford Avenue
P.O. Box 16268
Alexandria, VA 22302-0268

Dr. H. Wallace Sinaiko
Manpower Research
and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Paul Speckman
University of Missouri
Department of Statistics
Columbia, MO 65201

Dr. Martha Stocking
Educational Testing Service
Princeton, NJ 08541

Dr. Peter Stoloff
Center for Naal Analysis
200 North Beauregard Street
Alexandria, VA 22311

Dr. William Stout
University of Illinois
Department of Mathematics
Urbana, IL 61801

Maj. Bill Strickland
AF/MPXOA
4E168 Pentagon
Washington, DC 20330

Dr. Hariharan Swaminathan
Laboratory of Psychometric and
Evaluation Research
School of Education
University of Massachusetts
Amherst, MA 01003

Mr. Brad Sympson
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Kikumi Tatsuoka
CERL
252 Engineering Research
Laboratory
Urbana, IL 61801

Dr. Maurice Tatsuoka
220 Education Bldg
1310 S. Sixth St.
Champaign, IL 61820

Dr. David Thissen
Department of Psychology
University of Kansas
Lawrence, KS 66044

Mr. Gary Thomasson
University of Illinois
Educational Psychology
Champaign, IL 61820

Dr. Robert Tsutakawa
University of Missouri
Department of Statistics
222 Math. Sciences Bldg.
Columbia, MO 65211

Dr. Ledyard Tucker
University of Illinois
Department of Psychology
603 E. Daniel Street
Champaign, IL 61820

Dr. Vern W. Urry
Personnel R&D Center
Office of Personnel Management
1900 E. Street, NW
Washington, DC 20415

Dr. David Vale
Assessment Systems Corp.
2233 University Avenue
Suite 310
St. Paul, MN 55114

Dr. Frank Vicino
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Howard Wainer
Division of Psychological Studies
Educational Testing Service
Princeton, NJ 08541

Dr. Ming-Mei Wang
Lindquist Center
for Measuremet
University of Iowa
Iowa City, IA 52242

Dr. Thomas A. Warm
Coast Guard Institute
P. O. Substation 18
Oklahoma City, OK 73169

Dr. Brian Waters
Program Manager
Manpower Analysis Program
HumRRO
1100 S. Washington St.
Alexandria, VA 22314

Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

Dr. Ronald A. Weitzman
NPS, Code 54Wz
Monterey, CA 92152-6800

Major John Welsh
AFHRL/MOAN
Brooks AFB, TX 78223

Dr. Rand R. Wilcox
University of Southern
California
Department of Psychology
Los Angeles, CA 90007

German Military Representative
ATTN: Wolfgang Wildegrube
Streitkraefteamt
D-5300 Bonn 2
4000 Brandywine Street, NW
Washinton, DC 20016

Dr. Bruce Williams
Department of Educational
Psychology
University of Illinois
Urbana, IL 61801

Dr. Hilda Wing
Army Research Institute
5001 Eisenhower Ave.
Alexandria, VA 22333

Dr. Martin F. Wiskoff
Navy Personnel R & D Center
San Diego, CA 92152-6800

Mr. John H. Wolfe
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. George Wong
Biostatistics Laboratory
Memorial Sloan-Kettering
Cancer Center
1275 York Avenue
New York, NY 10021

Dr. Wendy Yen
CTB/McGraw Hill
Del Monte Research Park
Monterey, CA 93940