NAMRL- 1224

ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING

LT Dennis E. Egan, MSC, USNR (Ph.D)

27 February 1976

NAVAL AEROSPACE MEDICAL RESEARCH LABORATORY
PENSACOLA, FLORIDA

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NAMRL—1224
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING.
5. TYPE OF REPORT & PERIOD COVERED: Interim
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Dennis E. Egan, Ph.D. (LT, MSC, USNR)
8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Aerospace Medical Research Laboratory, Pensacola, Florida 32512
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: MF51.524.004-2011
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Medical Research Development Command, National Naval Medical Center, Bethesda, Maryland 20014
12. REPORT DATE: 27 Feb[redacted]76
13. NUMBER OF PAGES: 30
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): N/A
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)
18. SUPPLEMENTARY NOTES
19. KEY WORDS (Continue on reverse side if necessary and identify by block number):
Spatial ability. Computerized testing.
Response latency.
Response speed.
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The processing of visually presented spatial information is a critical component of the activities performed by pilots and aircrewmen. In particular, Radar Intercept Officers and Air Control Officers must make rapid and accurate spatial judgments. It is likely that variation in the ability to process spatial information accounts for some of the undesirable variations in the performance of these jobs. Previous research using conventional or "accuracy" scoring for paper-and-pencil tests has identified two "spatial factors" (Spatial Orientation and Spatial Visualization) that are valid predictors of success in pilot and navigator training programs. Recent experimental work has used the latency of response to spatial problems to analyze the mental processing of spatial information. The present studies combine these approaches by investigating both accuracy and latency scores as measures of the ability to process spatial information. Spatial test items were redesigned to be suitable for collecting latency as well as accuracy scores. In two

Item 20 (Continued)

experiments four new spatial tests were administered to groups of U. S. Navy pilot and Flight Officer Candidates. The psychometric properties of latency and accuracy scores from those tests were determined. Informal tests of several hypotheses about spatial processing were carried out. Derived measures of spatial processing were proposed and analyzed. Response latency scores are both feasible and desirable for assessing the ability process spatial information. Latency scores were highly reliable and correlated across different spatial tests. Accuracy scores were somewhat less reliable, but correlated predictably across tests. Interestingly, latency and accuracy were virtually independent measures. Tentative support was found for a model of Spatial Orientation patterned after theories of concept verification. Spatial Visualization appeared to be a continuous process similar to physically turning an object in space. Measures of spatial processing based on those models correlated in a consistent pattern.

Approved for public release; distribution unlimited.

ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING

Dennis E. Egan, Ph.D.
LT, MSC, USNR

Naval Air Systems Command
W 43-13 8881

Bureau of Medicine and Surgery
MF51.524.004 2011

Approved by	Released by
Ashton Graybiel, M. D. Assistant for Scientific Programs	Captain R. E. Mitchell, MSC, USN Commanding Officer

27 February 1976

Naval Aerospace Medical Research Laboratory
Pensacola, Florida 32512

SUMMARY PAGE

PROBLEM

The processing of visually presented spatial information is a critical component of the activities performed by pilots and aircrewmen. In particular, Radar Intercept Officers and Air Control Officers must make rapid and accurate spatial judgments. It is likely that variation in the ability to process spatial information accounts for some of the undesirable variations in the performance of these jobs.

Previous research using conventional or "accuracy" scoring for paper-and-pencil tests has identified two "spatial factors" (Spatial Orientation and Spatial Visualization) that are valid predictors of success in pilot and navigator training programs. Recent experimental work has used the latency of response to spatial problems to analyze the mental processing of spatial information. The present studies combine these approaches by investigating both accuracy and latency scores as measures of the ability to process spatial information. Spatial test items were redesigned to be suitable for collecting latency as well as accuracy scores. In two experiments four new spatial tests were administered to groups of U. S. Navy pilot and Flight Officer Candidates. The psychometric properties of latency and accuracy scores from those tests were determined. Informal tests of several hypotheses about spatial processing were carried out. Derived measures of spatial processing were proposed and analyzed.

FINDINGS

Response latency scores are both feasible and desirable for assessing the ability to process spatial information. Latency scores were highly reliable and correlated across different spatial tests. Accuracy scores were somewhat less reliable, but correlated predictably across tests. Interestingly, latency and accuracy were virtually independent measures. Tentative support was found for a model of Spatial Orientation patterned after theories of concept verification. Spatial Visualization appeared to be a continuous process similar to physically turning an object in space. Measures of spatial processing based on those models correlated in a consistent pattern.

ACKNOWLEDGEMENTS

The technical assistance of David Hannah, Steven Harris, and Peter Collyer in preparing programs and analyzing data is gratefully acknowledged.

INTRODUCTION

The processing of visually presented spatial information is a crucial part of many activities performed by pilots and aircrewmen. For pilots, these activities include the visual monitoring of the aircraft's position with respect to landmarks, horizon, or other aircraft. Aircrewmen such as Radar Intercept Officers (RIOs) and Air Control Officers (ACOs) must interpret electronically generated symbols to determine the relative position and speed of objects out of visible range. These and other important tasks requiring the processing spatial information are performed almost continuously while in flight. Consequently, the ability to process and use spatial information is an important predictor of success in the training of pilots and aircrewmen. The present studies are an attempt to gain further understanding of spatial information processing, and to explore the properties of new measures of spatial processes. The specific objectives of the experiments will be introduced following a brief review of studies of spatial information processing.

Background

Until recently, spatial information processing had been studied almost exclusively by applying factor analysis to batteries of paper-and-pencil, multiple-choice tests. Kelly (11) and Thurstone (15) were among the first to induce the existence of a "spatial factor." Since then, efforts have concentrated on isolating two or more spatial abilities through refinements in testing and statistical procedures. For example, Guilford and Lacey (7) were able to separate "Visualization" from "Spatial" ability and found evidence that the latter is composed of two distinct factors (labeled Space I and Space II). Visualization had high validity for predicting success in pilot training, and the Space I factor was a valid predictor of success in navigator and pilot training. In a review of the literature available at that time, French (4) identified a general Space factor (the ability to "perceive and compare spatial patterns") as well as two specific factors. Spatial Orientation was defined as the ability to "remain unconfused by varying orientations," while Visualization was described as the "comprehension of movements in a three-dimensional field."

Guilford (6) identified three spatial factors in his theory of the structure of intellect. These seem to represent current thinking about the factor structure of spatial abilities. The factors are: cognition of visual-figural systems (CFS-V), cognition of kinesthetic-figural systems (CFS-K), and cognition of figural transformations (CFT). Since the second of these factors (CSF-K) is specific to a single test, it will not be discussed further. The remaining two factors are well defined, each having been identified in 10 or more independent studies. Guilford's factors CFS-V and CFT show patterns of loadings quite similar to Space I and Visualization identified earlier (7). Thus, measures of CFS-V and CFT are valid for predicting success in pilot and navigator training.

Among the tests loading on CFS-V are the Guilford-Zimmerman Spatial Orientation (GZO) subtest (8), and Aerial Orientation, the predecessor of the Navy's Spatial Apperception Test (SAT). The Guilford-Zimmerman Spatial Visualization (GZV) subtest loads on the CFT factor. Recent work using these tests shows that they remain valid predictors of success in modern day pilot training. For example, Ambler and Smith (1) found that each test had a primary loading on a "Spatial Manipulation" factor in a study of aptitudes found in different aviation specialties. That factor was found to differentiate pilots from Naval Flight Officers, and to differentiate various specialties and achievement levels of pilots.

An alternative approach to the study of spatial information processing has been recently employed by Shepard and his colleagues (2, 12, 13). They have used the latency of response to individual items from tests of Spatial Visualization to analyze the mental processing of spatial information. For example, Shepard and Metzler (13) studied a task in which pictures of two three-dimensional block structures were presented and subjects had to decide whether the two figures were the same or different. Pictures of the same block could be presented at different orientations so that one figure had to be rotated through some angle to bring the two figures into physical congruence. The main finding was that the latency to make a correct "same" response was linearly related to the angle through which one figure had to be mentally rotated to bring it into congruity with the other figure. Snyder (14) explored derived measures of performance for individual subjects on this task, and found systematic relationships between these measures and scores on tests of spatial and imagery abilities. Shepard and Feng (12) demonstrated a linear relationship between complexity of a mental paper-folding task and the latency of response. Cooper and Shepard (2) gave an excellent review of this work, and pursued several theoretical questions in a series of experiments.

These findings suggest that the mental processing that occurs in tests of Spatial Visualization is continuous in real time. While speeded paper-and-pencil tests of Visualization may reflect the time-to-process information, accuracy scores on such tests are less direct measures than the actual processing times. However, there may be information contained in accuracy but not in latency, so that the two types of measures together may provide a more complete assessment than would either taken alone. As shown in the following, Spatial Orientation may also be considered a real-time process, so that processing times may be desirable measures for that ability as well. Finally, it should be noted that the binary-choice format used by Shepard and his colleagues minimizes the impact of answer elimination strategies peculiar to multiple-choice tests. Thus, a "Yes"/"No" response format allows for more precise measurement of spatial processing.

Findings relating the ability to process spatial information to success in aviation can be summarized as follows. First, using accuracy scores, the existence of at least one "spatial factor" has been firmly established. It is probable that more than one factor can be identified. Second, certain tests loading on these spatial factors have proved to be valid predictors of success in aviation. Third, new measures based on latency of response to spatial problems may more precisely capture the mental processing of spatial information. This technique may be especially powerful if test items are designed for binary responses.

The above summary suggests that an investigation of the latency of response to spatial items may yield better measures for predicting pilot or aircrewman success in training. There are two additional reasons for studying the time taken to process spatial information. One is that certain spatial tasks in aviation are time critical. In particular, the RIO must respond to displays presenting rapidly evolving spatial information. In such cases speed as well as accuracy is required, and a measure of speed of processing spatial information may be a valid predictor of performance. A second reason for studying latency is that some available data on non-spatial tasks (10) suggests that latencies can be reliable yet virtually independent of accuracy scores. Potentially, the speed of response may yield information about a candidate that is not contained in the traditional measure of accuracy.

Objectives

The tasks of RIOs, ACOs and other Naval Flight Officers impose a heavy requirement for the processing of spatial information. Although each of these specialties involves an intensive and highly technical training program, individuals can still be identified who are considered deficient in some critical job aspects. Further, these deficiencies are often not remediable by additional training of the same type. It is probable that variations in the spatial abilities of the operators can account for some of these undesirable variations in job performance. If these abilities can be identified, and their role in performing the tasks can be determined, remedies for deficiencies in ability can be achieved through selecting or through more appropriate training procedures. The experiments reported here are an initial step in defining and assessing the spatial abilities present in the naval aviation community. The studies are organized around the following specific objectives.

1. Select several spatial tests and redesign items in a way that allows for collecting latency and accuracy of responses in a group testing situation. The tests should include representatives from the two major categories of spatial tests, Spatial Orientation and Spatial Visualization.

2. Obtain the psychometric properties of accuracy and latency for the new tests. In addition to means and variances, the reliability of any measure should be examined.

3. Obtain intercorrelations of scores. The pattern of correlations most desirable for the purpose of developing new measures would have the following characteristics. (i) Accuracy scores across all tests should correlate significantly. Ideally, correlations between paper-and-pencil forms and redesigned

forms of the same test should be at or near the level of alternate-form reliability. Correlations among accuracy scores on different spatial tests purported to measure the same factor should be slightly lower. Correlations of accuracy scores on tests of different spatial factors should be lower yet, but still significant. This pattern is known to occur for the paper-and-pencil forms of spatial tests. These findings would indicate that the redesigned tests are measuring the same quantity as their paper-and-pencil counterparts, and that different spatial tests are measuring common spatial processes. (ii) Latency scores on different tests should correlate significantly. This feature would indicate that latency is measuring processes common to all of the redesigned tests. These correlations should be highest for tests measuring the same spatial process or factor. (iii) Latency and accuracy scores should be independent. If latency scores are to be of use, they should yield information about spatial processes that is not contained in accuracy scores. Specifically, both high positive accuracy-latency correlations (speed-accuracy tradeoff) and high negative accuracy-latency correlations (measurement of the same phenomenon) are undesirable.

4. Propose models of spatial information processing and test those models. This objective aims to extend the theoretical work of Shepard and his colleagues, and to develop a theory of Spatial Orientation. While rigorous model testing cannot be accomplished in experiments designed mainly for establishing the characteristics of new measurement instruments, certain informal tests will be carried out.

5. On the basis of the models, propose and analyze derived measures of spatial information processing. These derived measures may give the most precise estimates of the ability to process spatial information. If so, they may be very useful in predicting criteria such as RIO air intercept performance.

EXPERIMENT I

The first experiment drew items from the Navy's current SAT, and the GZV and GZO subtests.

METHOD

Subjects

The examinees were 30 Aviation Officer Candidates (AOCs) and 32 Naval Flight Officer Candidates (NFOCs) who were available for testing during their first week of indoctrination at Pensacola Naval Air Station. Because of scheduling difficulties and equipment failures, complete data were available for only 31 examinees. Each examinee had been selected by the Navy for admission into the AOC or NFOC program on the basis of a battery of screening tests. A major component in that battery is the SAT. Consequently, typical examinees in this study had greater spatial ability than average applicants who were college graduates.

Apparatus

Construction of Test Items. The new version of the SAT designed for latency scoring (LSAT) was constructed from multiple-choice items from Form A and Form B of the SAT. The LSAT requires the examinee to judge whether a landscape shown in one panel is the view that would be seen from the cockpit of an airplane shown in another panel. The standard SAT presents for each of 30 landscapes a set of five airplanes shown at different orientations. An item from each test is given in Figure 1.

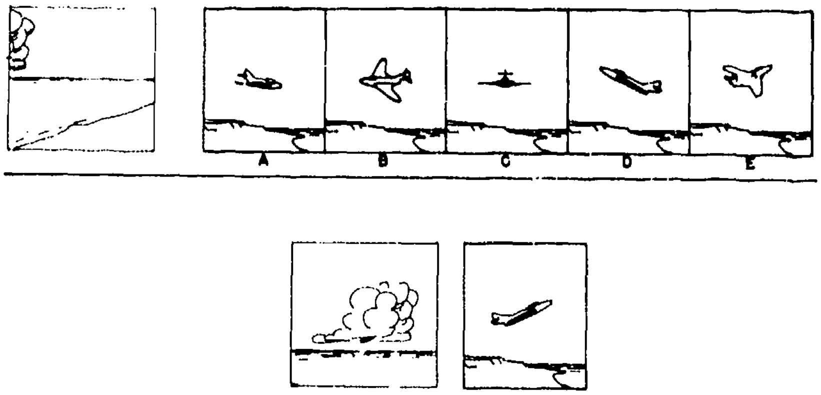

Figure 1. An Item From the Spatial Apperception Test (top) And an Item from the Redesigned Test (bottom)

In the SAT the examinee selects the best choice for each item and has a time limit of 10 minutes for the entire test. In the LSAT, examinees had a maximum of 15 seconds per item to make a "Yes" or "No" response. The 60 items for the LSAT were inter-leaved in order from the two forms of the SAT (30 items from each) so that item k in either form of the SAT appeared randomly in position 2k or 2k-1 in the LSAT. Half of the items were randomly selected to be "Yes" items, and the other half were "No" items. For "Yes" items the landscape was matched with the correct airplane from the SAT. For "No" items, the landscape was paired with a randomly selected false choice.

The LGZV was constructed in a similar manner from the 40-item multiple-choice GZV (Form B)*. The GZV requires examinees to mentally manipulate an alarm clock according to a specified sequence of rotations and then to judge which of five figures matches its final position. Each item of the GZV consists of one view of an alarm clock, a figure depicting the required rotations, and a set of five clocks shown in different final orientations. An item from the GZV and one from the LGZV are shown in Figure 2.

---

*Permission was obtained from Sheridan Psychological Services, Inc., to use the GZV and GZO subtests in these studies.

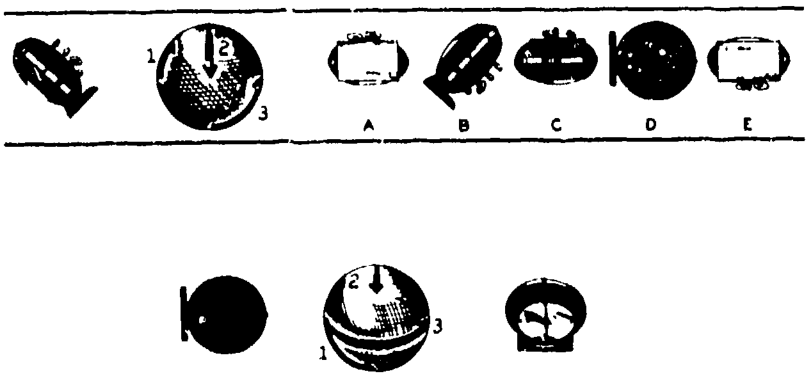

Figure 2. An Item from the Guilford Zimmerman Spatial Visualization Test (top) and an Item From the Redesigned test (bottom)

In the GZV the examinee selects the best choice for each item and has a time limit of 10 minutes for the entire 40-item test. In the LGZV examinees were given a maximum of 20 seconds per item to make a "Yes" or "No" response. Items in the LGZV were presented in the same order as they occurred in the GZV. This orders items by their difficulty, since those requiring more rotations are presented later in the test. True and false choices were randomly determined as in the LSAT.

The third spatial test, the LGZO, was constructed from Form A of the GZO. This test requires examinees to determine whether a symbol accurately portrays the change in position and direction that has occurred from the top to the bottom drawing of a motorboat heading toward a coastline. The GZO presents 60 items consisting of the two drawings and a set of five symbols. An item from the GZO and one from the LGZO are shown in Figure 3.

In the GZO the examinee selects the symbol that best portrays the change that has occurred from the top to the bottom picture. The time limit on the 60-item test is 10 minutes. In the LGZO, examinees were given a maximum of 15 seconds to respond "Yes" or "No" to each item. The order of presentation was the same in the two tests, and selection of true and false items in the LGZO was again determined randomly.

Instructions for the three redesigned tests were simple modifications of the instructions for the paper-and-pencil forms. The modified instructions showed examples of the items and explained the use of the testing apparatus. They also included a statement to be as accurate as possible, and informed the examinees of the maximum time limit allowed for each item.

<u>Test Apparatus</u>. The new tests were given on the Multiple Unit Test System at NAMRL. The system controlled the presentation, timing, and scoring of two-choice test items for groups of six or fewer examinees. The system comprised

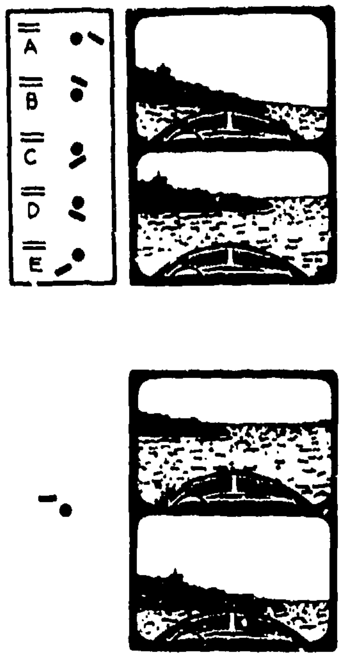

Figure 3. An Item from the Guilford Zimmerman Spatial Orientation Test (top), and an Item From the Redesigned Test (bottom)

six testing stations, a Kodak Ectagraphic self-focusing slide projector (Model AF-2), and a centrally located viewing screen. A UNIVAC 418 computer operating in the real-time mode controlled the system. Test stations were arranged in a row parallel to the screen and between the screen and projector. The screen was 4.24 meters in front of the stations. The row of stations was placed so that the viewing angle at the two outboard stations was no larger than 30°. Each station was equipped with a hand-held switchbox on which two response buttons were mounted. The lefthand button was labeled "No" and the righthand button was labeled "Yes." Examinees were instructed to hold the box in their hands and use their thumbs to activate the buttons.

Procedure. Examinees were given the new tests in the following order: LSAT, LGZV, LGZO. All examinees took the LSAT, and depending on their schedules and the availability of equipment subsequently received the LGZV then the LGZO. The smaller number of examinees taking the LGZO was thus a proper subset of those taking the LGZV, etc. Three to five days after taking the new versions, examinees were given the standard versions of the GZV and GZO. These paper-and-pencil forms were administered under group testing conditions with approximately 25 examinees per group. The SAT had been given prior to admission to the program, so those scores were obtained from the examinee's records.

The procedure for the new tests started with the examinees reading the instruction booklet. When all indicated they understood the instructions, the experimenter gave a verbal "Ready" signal and initiated the test. Items were projected onto the screen and examinees responded by pushing the appropriate button on the switchbox. An item remained in view on the screen until either all examinees responded, or the maximum time limit was reached. Approximately 1.5 seconds after the first of those events occurred, the slide projector advanced to the next item. After a succession of six such items, a blank trial was presented and allowed to time out. This served as a short rest. Just prior to the initiation of the next sequence of six items, a "Ready" signal was given.

Scoring. Latency of response to an item was defined as the interval between the onset of presentation and the completion of the response to the item. Latencies and answers were stored by the computer at the time of testing and later transferred to magnetic tapes for data reduction and analysis. If an examinee did not answer an item by the end of the time limit, the item was scored as wrong, with a latency equal to the time limit. Latencies for wrong answers were not used except for tests of the Visualization model (cf. Fig. 6).

RESULTS

Psychometric Properties

The mean, standard deviation, and reliability of accuracy scores on the six tests of spatial ability are given in Table I. Also included in Table I are the properties of the latency of response for items on the LSAT, LGZV, and LGZO. These data are not set forth as norms, since the absolute values of scores, especially latency scores, undoubtedly depend on the design and calibration of the testing apparatus. However, these properties do permit several useful observations.

Table I

Means, Standard Deviations and Reliabilities of Accuracy and Latency Scores (Experiment I)

Measure	N	Mean	S.D.	Reliability
LSAT Number Correct	61	45.51	6.41	.81[a]
LSAT Mean Correct Latency (sec.)	61	5.8d	1.13	.93[a]
LGZV Number Correct	52	26.02	5.26	.80[a]
LGZV Mean Correct Latency (sec.)	52	10.51	1.67	.89[a]
LGZO Number Correct	32	44.84	8.04	.93[a]
LGZO Mean Correct Latency (sec.)	32	7.05	1.16	.92[a]
SAT Number Correct	61	19.82	5.77	.71[b]
GZV Number Correct	54	24.93	8.06	.91[c]
GZO Number Correct	54	32.65	11.64	.88[d]

[a]Split-half (odd/even) reliability corrected for test length.
[b]Uncorrected alternate-form reliability (5).
[c]Split-half reliability reported in Guilford & Zimmerman (9).
[d]Reliability estimated by administering test in two separately timed, equivalent halves, intercorrelating the part scores, and applying the Spearman-Brown formula (9).

The percentage of items answered correctly was generally greater on the new versions of the tests, since a guess will result in a correct answer more often in a two-choice than a five-choice problem. Also, examinees will at least attempt each problem in the new tests, whereas in the standard versions some items occurring later in a test may never be attempted.

The LGZV was the most difficult test in the entire battery. If corrected for guessing, scores on it would be considerably below any of the other tests. The mean latency for correct responses to LGZV items was the highest, and the time limit was exceeded on a greater proportion of items from the LGZV (0.047) than either the LGZO (0.017) or the LSAT (0.007).

In the case of the LSAT and LGZV, mean latency was substantially more reliable than accuracy. For the LGZO, reliabilities of accuracy and latency were about the same. The lower reliability for accuracy is again explained by the high probability of a guess being correct in two-choice tests. The two-choice format resulted in higher means and smaller standard deviations for tests of a given length. Reliabilities of latencies approximate those of accuracy scores on the standard tests, except for the LSAT for which the split-half reliability of latency exceeds the alternate-form reliability of the standard SAT score.

Intercorrelations

The correlations among latency and accuracy scores are shown in Table II. The Ns for these correlations vary from 31 to 61 depending on available data. The pattern of correlation shows several desirable characteristics. First, correlations among accuracy scores on all the tests were generally statistically significant. The highest correlations between accuracy scores occurred when comparing accuracy on the standard and redesigned forms of the same test. The correlations between the GZV and LGZV ($r = .74$) and between the GZO and LGZO ($r = .72$) are satisfactorily close to alternate-form reliability when restriction of range of ability in the sample is considered. The correlation of accuracy scores on the LSAT and SAT was lower ($r = .53$) but still highly significant. This lower correlation is probably due to the fact that the SAT scores were derived from two different forms of the test administered many months before the examinees participated in the experiment. The difference between Spatial Orientation and Spatial Visualization was not apparent in these data since accuracy scores from different tests of the same factor did not correlate at a higher level than accuracy scores from tests of different factors. Excluding correlations between a test and its redesigned version, the mean of correlations among tests of Spatial Orientation (LSAT, SAT, LGZO, GZO) was actually slightly lower ($\bar{r} = .39$) than the mean of correlations between tests of Orientation and Visualization ($\bar{r} = .45$). Generally, these data indicate that the accuracy scores on all tests measured a common process or ability, and that the distinction between factors of Orientation and Visualization was not reflected in the accuracy scores.

A second characteristic of the data in Table II is that the three measures of latency were highly correlated. Thus the mean time taken to respond correctly to spatial test items was a consistent characteristic of an examinee across all three redesigned tests.

Table II

Correlations of Accuracy and Latency Scores[a]
(Experiment I)

		2	3	4	5	6	7	8	9
1.	LSAT Number Correct	-.15	.40**	-.46**	.39*	-.40*	.53**	.46**	.49**
2.	LSAT Mean Correct Latency		-.04	.57**	-.14	.58**	-.14	-.31	-.23
3.	LGZV Number Correct			-.30*	.30	-.10	.40**	.74**	.53**
4.	LGZV Mean Correct Latency				-.43*	.46**	-.39**	-.46**	-.57**
5.	LGZO Number Correct					-.27	.26	.47**	.72**
6.	LGZO Mean Correct Latency						.00	-.34	-.30
7.	SAT Number Correct							.45**	.43*
8.	GZV Number Correct								.61**
9.	GZO Number Correct								

**$p \lesssim .01$
*$p < .05$
[a]Sample sizes range from 31 to 61.

Third, correlations between latency and accuracy scores were generally negative and of low magnitude. The main exception was the latency score on the LGZV that correlated significantly negative with each measure of accuracy. Explanations for this exception can be advanced, but the result should first be replicated. The low correlations between accuracy and latency suggests that, for the conditions studied, speed and accuracy of spatial processing are for practical purposes independent. Whether the speed being measured is peculiar to spatial processing, or whether it is a more general personality or intelligence factor, cannot be determined by these results.

SUMMARY

Experiment I demonstrated that spatial tests can be designed to yield accuracy and latency scores that are reliable and have a desirable pattern of correlation. A second experiment was performed to replicate those findings.

EXPERIMENT II

The main difference between the first and second experiments was that the LGZO was replaced by a block rotation test, LBRT, in the battery. The new test used items similar to those employed by Shepard and Metzler (12) in their study of mental rotation. Standard forms of the test show loadings on the Spatial Visualization factor (4).

METHOD

Subjects

The examinees were 28 AOCs, 23 NFOCs, 19 Aviation Reserve Officer Cadets, and 2 Air Intelligence Officer Candidates. Six examinees were tested in each session. Because of scheduling and equipment difficulties, complete data were available for only 48 examinees.

Apparatus

The LSAT and LGZV were identical to those used in Experiment I. A block rotation test (LBRT) was constructed. For this test, three rigid three-dimensional block structures were drawn. Photographs of the drawings and their mirror images were taken. Each test item consisted of two of these figures. The pair was either the same block structure presented in two different orientations, or one figure and its mirror image. Three sets of items were constructed, one for each block figure. In each set, 9 items presented a pair of identical figures at varying orientations, and 9 items presented a figure and its mirror image. The match items were constructed so that the difference in angular orientation of the two figures was an integer multiple of 40°. Therefore, rotation in the vertical plane of 0°, 40°, 80°, etc., was required to bring the two figures into congruence. For purposes of analysis, figures differing by k degrees left or right of zero were grouped together. The nine match figures in each set thus differed by 0°, $\pm 40^{\circ}$, $\pm 80^{\circ}$, $\pm 120^{\circ}$, or $\pm 160^{\circ}$. The total number of items was 54, 9 match and 9 no-match items for each of three basic figures. A match and a no-match are shown in Figure 4. Items were arranged randomly in blocks of six as in the first experiment. The order of items was the same for each examinee. The test apparatus was the same one used in Experiment I.

Procedure

The procedure was identical to that in the first experiment except that the LBRT was substituted for the LGZO. Scoring was the same as in Experiment I.

RESULTS

Psychometric Properties

Means, standard deviations, and reliabilities for latency and accuracy scores are given in Table III. Across the two experiments, the psychometric properties of the LSAT were quite similar (see Table I). Performance on the LGZV in Experiment II was at a higher level and less variable. The LBRT proved to be the easiest test with the consequence that reliability of accuracy scores was

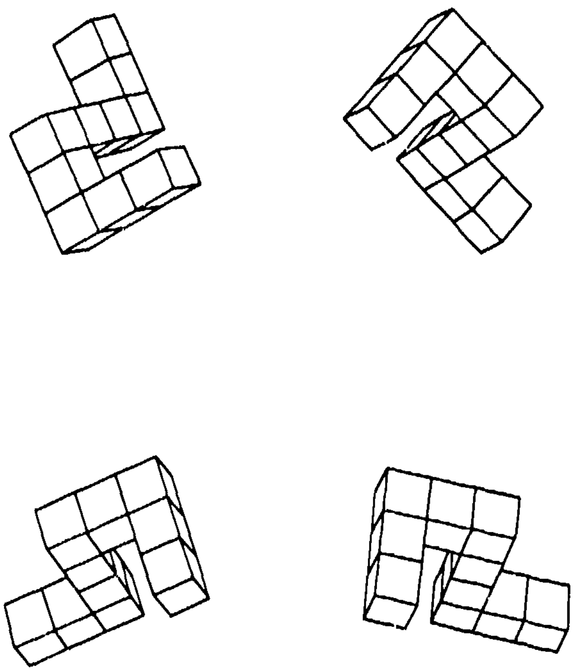

Figure 4. A "YES" Item (top) and a "NO" Item (bottom) From the Block Rotation Test

not high (.65). However, reliability of mean latency to respond correctly on the LBRT (.92) was acceptable. In all cases, reliability of a latency was substantially higher than the corresponding accuracy reliability. That result is probably due in part to the two-choice format for test items.

Table III

Means, Standard Deviations and Reliabilities of Accuracy and Latency Scores

(Experiment II)

Measure	N	Mean	S.D.	Reliability[a]
LSAT Number Correct	66	45.14	6.84	.76
LSAT Mean Correct Latency (sec.)	66	6.14	1.16	.95
LGZV Number Correct	54	27.35	4.58	.62
LGZV Mean Correct Latency (sec.)	54	10.35	1.32	.78
LBRT Number Correct	60	45.85	4.75	.65
LBRT Mean Correct Latency (sec.)	60	6.39	1.25	.92

[a]Split-half (odd/even) reliability corrected for length of test.

Intercorrelations

Table IV gives the intercorrelation of accuracy and latency scores. The Ns for these correlations vary from 48 to 66. The pattern of correlation was similar to that found in Experiment I. Correlations among accuracy scores were positive and high as were correlations among latency scores. Again, however, accuracy-latency correlations tended to be negative and low in magnitude. The exception to this pattern in Experiment I was the LGZV mean latency score which was significantly correlated with all other scores. That result was not replicated in Experiment II, but the general patterns of correlations in the two studies were very similar.

Table IV

Correlations of Accuracy and Latency Scores[a]
(Experiment II)

		2	3	4	5	6
1.	LSAT Number Correct	-.12	.49**	.00	.40**	-.11
2.	LSAT Mean Correct Latency		-.12	.39**	-.02	.45**
3.	LGZV Number Correct			-.18	.46**	-.19
4.	LGZV Mean Correct Latency				-.25	.75**
5.	LBRT Number Correct					-.26*
6.	LBRT Mean Correct Latency					

** $p \leq .01$
* $p < .05$
[a]Sample sizes varied from 48 to 72.

THEORETICAL IMPLICATIONS

The foregoing empirical results show that accuracy and mean latency of responses to spatial problems are desirable measures of the ability to process spatial information. At this point, information-processing analyses of the experimental tests will be introduced for two reasons. First, the theoretical development should lead to a greater understanding of the fine structure of making a response to a spatial problem. This understanding could prove useful when conducting analyses of criterion tasks (e.g., those performed by RIOs and ACOs). Second, information-processing analyses should suggest additional measures that are direct estimates of theoretical parameters. Estimates of these parameters for individual subjects may be the most precise measures of spatial ability.

Information Processing Analysis: Visualization

A simple model for the mental clock-turning task required by the LGZV is depicted in Fig. 5. The model is based on the assumption that each item requires the examinee to store visual information, then perform a sequence of mental rotations, then compare the result with the figure given, and finally respond. The middle part of this process is a loop performed once for each rotation required. The inpu., match, and response stages are performed only once regardless of the number of turns required.

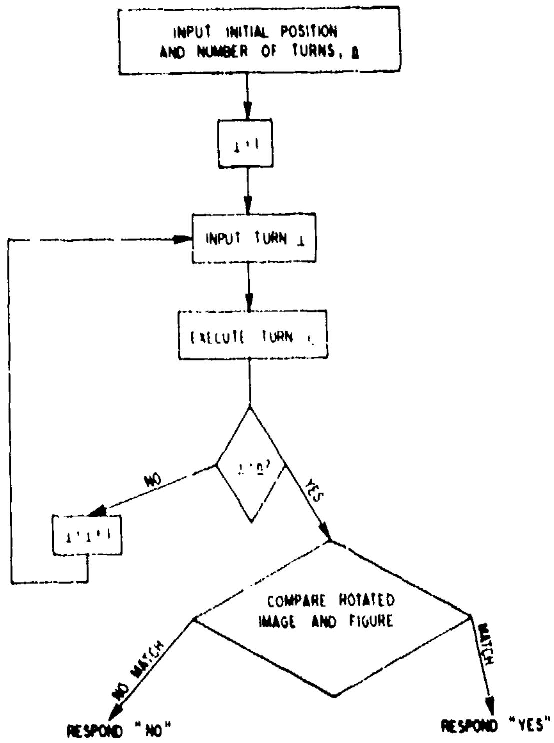

Figure 5 Information Processing Model of Spatial Visualization Task

Given a hypothetical model of the spatial visualization task, two kinds of analyses are required. One is to test the model to determine whether it accurately characterizes human performance on visualization tasks. Given evidence that the model predicts performance, the second analysis is to select dependent measures from the task that capture the performance of individuals.

The tests of the proposed models are limited by the experimental method employed. Typically such tests would be conducted with appropriate experimental controls and randomizing procedures. Since the emphasis of the present studies was to develop measures of individual differences, the usual experimental procedures were not practical to use. Thus, the evaluation of the model in Figure 5 and subsequent models should be considered tentative until further experimental work is done.

Two predictions derived from the visualization model will be tested. The first is that both latency of response and error rate should increase with the number of turns required by the LGZV item. This prediction is based on the fact that the input-rotate-decide sequence must be performed once for each turn. Additional time and complications are involved as more sequences are required. A stronger prediction is possible if it is assumed that (i) the input-rotate-decide loop will take a constant amount of time, $\underline{t}$, for each cycle, and (ii) that the initial input stage and final match and output stages take a constant time, $\underline{k}$, regardless of the number of turns required. Under these assumptions, response latency is predicted to be a linear function of the number of turns required, $\underline{n}$. Mean latency for an item requiring $\underline{n}$ turns is given by $L_n = \underline{k} + \underline{nt}$.

The second prediction is based on the assumed locus of error in the process shown in Figure 5. It seems unlikely that an examinee would execute the wrong number of turns for an item, because that number is clearly indicated on the slide. Given that $\underline{n}$ turns are performed, latency of response will be directly related to $\underline{n}$ whether the answer is an error or not. Thus the second prediction is that latency for correct and wrong answers will follow about the same pattern. The two predictions can be evaluated by the data in Figure 6 which gives the mean latency of correct and wrong answers (Experiment I) for items requiring $\underline{n}$ turns. Data from Experiment II followed the same pattern.

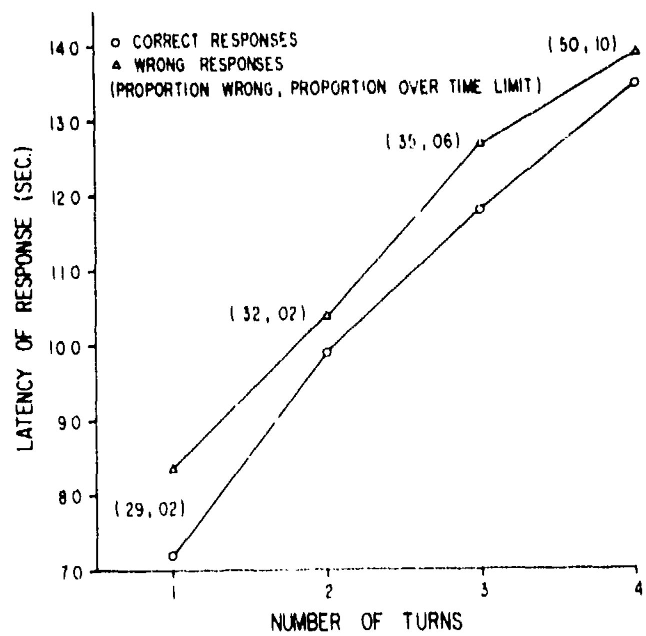

Figure 6. Mean Latency for Correct and Wrong Answers to Items Requiring n Turns on the [illegible] (Experiment I)

Both predictions were well supported. Latency of response for correct and wrong answers monotonically increased over number of turns required. The relationship between latency and turns appears approximately linear. Error rates and percentage of responses exceeding the deadline increased in a monotonic fashion with number of turns required. These observations lend credibility to the Visualization model.

Given tentative support for the model, the obvious measures for comparing individuals are the slope and intercept of an examinee's curve relating response latency to number of turns. The zero-intercept, k, captures the amount of time taken to store the visual stimulus, check the rotated mental image against a visual pattern, and respond. The slope, t, gives the rate at which the input-rotate-decide cycle can be performed. These parameters were calculated for each examinee. Latencies for correct and wrong responses were pooled to obtain more reliable estimates of the slopes and intercepts. Of the 106 examinees taking the LGZV in the two studies, only 1 had a negative slope, and 1 had a negative intercept. Thus the pattern shown in Figure 6 was true not only for grouped data, but also for the majority of individual examinees.

Based on the findings of Shepard and his colleagues, predictions can be advanced for the other test of Visualization, the LBRT. First, latency and error rates of "Yes" items should be directly related to the angular difference between the two block figures. Shepard and Metzler (1971) found a linear relationship between latency and angular difference in highly practiced subjects responding to 1600 randomly ordered stimuli. For unpracticed examinees attempting 54 problems presented in a fixed order, at least an increasing monotonic function should be obtained. Second, the mean latency of correct "No" responses is predicted to be greater than the mean latency of correct "Yes" responses. This prediction is derived from the idea that a "No" response is made only after all mental rotations have been tried, but a "Yes" response may occur after a varying amount of rotation depending on the angular orientation of the two figures. The same idea motivates the third prediction, that the variability of correct "Yes" responses should be greater than that of correct "No" responses. These predictions are evaluated in Figure 7 and Table V.

Table V

Means and Standard Deviations of Latencies for Correct "Yes" and "No" Responses

	LBRT	LSAT (Exp. I)	LSAT (Exp. II)	LGZO
"Yes" Items				
Mean	5.98 sec.	6.06 sec.	6.47 sec.	6.99 sec.
S.D.	1.70	0.60	0.70	0.90
"No" Items				
Mean	7.16 sec.	5.80 sec.	6.07 sec.	7.08 sec.
S.D.	1.57	0.81	0.78	0.76

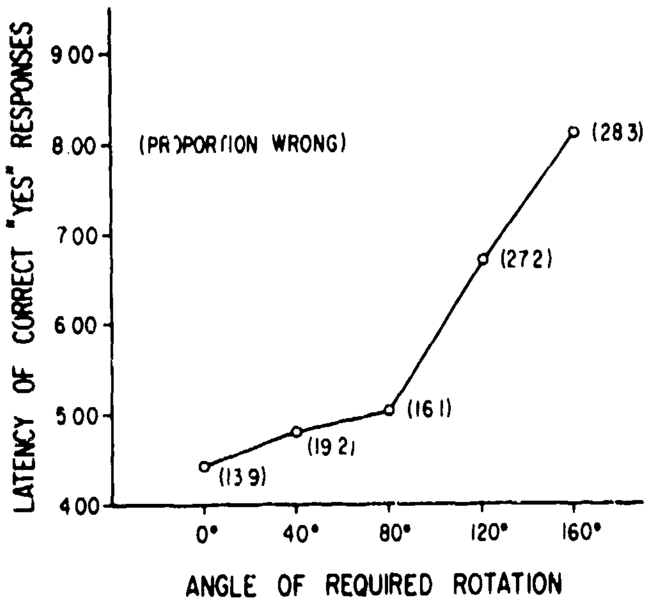

Figure 7 Mean Latency for Correct "YES" Responses as a Function of Required Angle of Rotation to Achieve a Match on the LBRT

The data in Figure 7 indicate that both latency of response and error rate were related to angular difference between stimulus figures in a monotonic fashion. As shown in Table V, mean latency of correct "Yes" items was less than mean latency for correct "No" items. Furthermore, there was greater variability among means of "Yes" items than means of "No" items. Thus all three predictions received tentative support from these data. For each examinee the best-fitting line relating latency of correct "Yes" responses to angular difference was computed. The slopes and intercepts of these lines were used as additional measures of spatial processing.

Estimates of individual slopes and intercepts tended to support the idea that these parameters of a straight line capture the actual mental processing that occurs in Spatial Visualization. Of the 60 slopes calculated for the LBRT, only 1 was found to be negative. Of the 60 intercepts, all were positive. In summary, for the tests of Visualization used here, larger angles of a single mental rotation, and greater numbers of rotations resulted in longer latencies of response in data for groups and individuals.

<u>Information Processing Analysis: Orientation</u>

One way to conceive of Spatial Orientation is to consider it a form of concept verification (3). On this view, an examinee serially selects and tests the three spatial dimensions of a visual pattern against his concept of what that pattern ought to be. Accordingly, a hypothetical description of the Spatial Orientation process in these experiments is given by the decision tree in Figure 8.

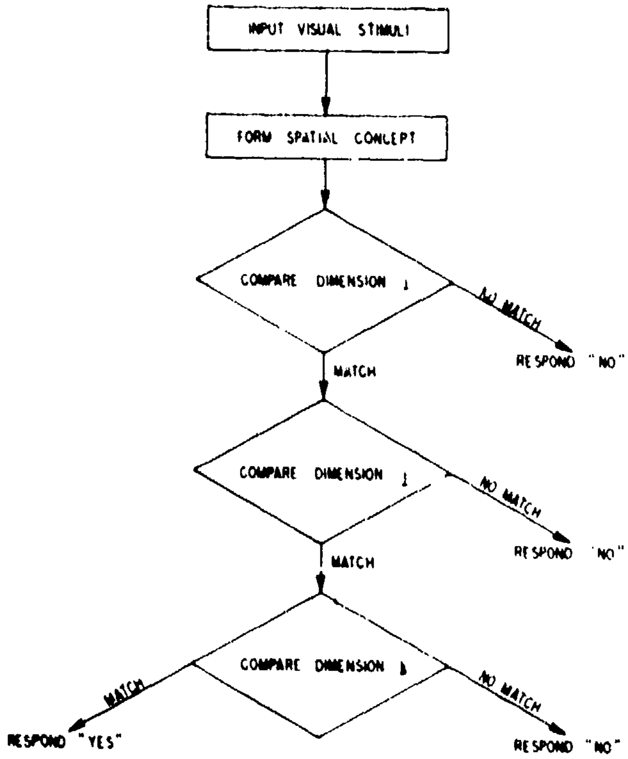

Figure 8. Information Processing Model of Spatial Orientation Task

The Model in Figure 8 will be used to generate predictions for the LSAT and the LGZO. For each of those tests, an item is wrong only if the test figure shows one or more discrepancies from the correct spatial concept. Discrepancies can occur in any of three dimensions: heading, pitch, or bank. The model assumes that the examinee checks the orientation of the airplane in the LSAT and the bar symbol in the LGZO along each of these dimensions. If a check of the first dimension reveals that the test figure does not correspond to the correct concept, a "No" response is given. If the figure matches the correct concept on the first dimension, checks of each of the two remaining dimensions must be made before a "Yes" response is given. The model thus tests a relational structure (3) that can be described as a three-dimensional conjunctive concept. Values on each of the three dimensions are checked for error, and "Yes" responses are possible only when all dimensions are found to match the concept.

Three predictions will be derived from the model in Figure 8 and applied to data from the LSAT and LGZO. The first prediction is that response latency and error rate should be inversely related to the number of dimensions on which the test figure differs from the correct spatial concept. If a difference exists on all three dimensions, it will be detected when the first dimension is selected and compared. If a difference exists on fewer dimensions, one or more dimensions may result in a correct match before the discrepancy is found. This is predicted to result in (i) longer latency of a correct "No" response as more dimensions are checked, and (ii) greater likelihood of error caused by examinees failing to detect a difference or failing to test all dimensions and guessing.

The second prediction is that the mean latency of correct "Yes" responses will be greater than the mean latency of correct "No" responses. This prediction is derived from the fact that correct "No" responses require fewer comparisons than correct "Yes" responses (ignoring guessing). The latter can occur only after all three dimensions have been tested.

The third prediction is derived from the fact that correct "No" responses can occur after a variable number of comparisons, but correct "Yes" responses (unless a result of a guess) must be based on the outcome of exactly three tests. The prediction is that the variation of mean latencies for "No" items will be greater than the variation of latencies for "Yes" items. These three predictions are evaluated for the LSAT and LGZO in Figure 9 (data from Experiment I) and Table V (data from Experiments I and II).

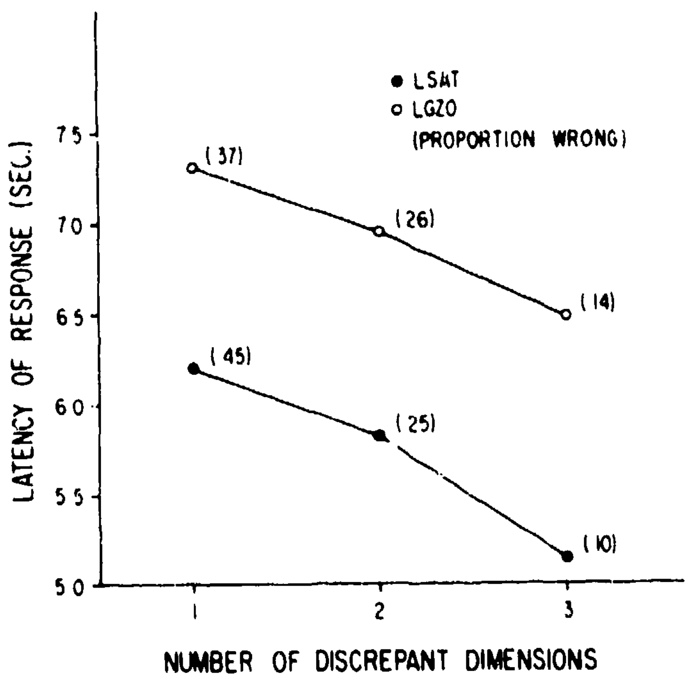

Figure 9 Mean Latency for Correct "NO" Responses on the LSAT and LGZO as a Function of the Number of Discrepancies Between the Correct Spatial Concept and the Figure Presented (Experiment I)

Generally, the predictions and data correspond well for the LSAT, but to a lesser degree for the LGZO. In particular, latency and error rates of correct "No" responses were inversely related to the number of dimensions on which items were discrepant from the correct spatial concept (see figure 9). This pattern was true for the LSAT in both experiments and LGZO in Experiment I. As predicted, correct "Yes" responses had greater mean latency and less variability than correct "No" responses for the LSAT. That prediction did not hold for the LGZO. The impression given by the present data is that some support exists for the serial decision model, especially as applied to the LSAT.

Regression lines relating latency of correct "No" responses to the number of discrepant dimensions were calculated for each examinee on both the LSAT and LGZO. The slope and zero-intercept of each line were selected as measures best characterizing the spatial orientation process. Intercepts were viewed as estimating the sum of time taken to input, conceptualize, and respond to spatial orientation problems. The slopes of the lines were taken as a measure of the speed with which an examinee could select and match on each additional spatial dimension. Agreement between data from individuals and the model was not as great as that found for Spatial Visualization. While no negative intercepts were observed, for the LSAT 28 of 127 examinees yielded positive slopes contrary to expectation. For the LGZO, 8 of 32 examinees had positive slopes.

Analysis of Derived Measures

Correlations involving slopes and intercepts on each of the new spatial tests are given in Table VI. The highest correlations occurred between the slopes and intercepts derived from the same test. In every case, examinees characterized by steep slopes; i.e., those taking a long time to make additional decisions or rotations, tended to have smaller intercepts which are assumed to estimate time for input and output processes. Slopes did not correlate significantly with other measures as often (6 correlations of a possible 50 reached statistical significance) as did intercepts (26 of 50 significant). Correlations between intercepts and mean response times were all positive with a mean of $\bar{r} = .39$. Correlations between intercepts and number correct were generally negative and had a mean of $\bar{r} = -.19$.

Concerning derived measures on the two tests of visualization (LGZV and LBRT), intercepts correlated significantly ($r = .40$, $p < .01$) as did slopes ($r = .27$, $p < .05$). This expected pattern was not found for derived measures on tests of orientation (LSAT and LGZO).

DISCUSSION

In the two experiments, consistent evidence was obtained regarding each objective of this research. The feasibility of collecting and interpreting accuracy and latency data in tests of Spatial Orientation and Visualization has been demonstrated. Further research should investigate possible effects due to viewing

Table VI

Correlations Involving Derived Measures of Spatial Information Processing[a]

		12	13	14	15	16	17	18	19
1.	SAT Number Correct.	-.22*	.31**	-.06	-.09	.32	.19	-.05	.01
2.	LSAT Number Correct.	-.06	-.22*	.16	-.28**	.29	-.10	.22	-.33*
3.	LSAT Mean Correct RT.	-.21*	.49**	-.12	.42**	.29	.62**	.04	.24
4.	GZV Number Correct.	-.04	-.22*	.20	-.41**	.03	-.24	-.06	-.31*
5.	LGZV Number Correct	-.06	-.16	.04	-.22*	.00	-.08	-.01	-.13
6.	LGZV Mean Correct RT	-.05	.24*	-.13	.66**	-.24	.17	.08	.32*
7.	GZO Number Correct	-.08	-.21*	.34**	-.50**	.29	-.05	.27	-.36*
8.	LGZO Number Correct	-.24	-.36*	.27	-.51**	.24	-.07	. .	. .
9.	LGZO Mean Correct RT	-.23	.15	.06	.35	-.03	.65**	. .	. .
10.	LBRT Number Correct	.04	-.03	-.02	-.07	. .	. .	-.01	-.20
11.	LBRT Mean Correct RT	-.27*	.06	.19	.61**	. .	. .	.24	.43**
12.	LSAT Slope		.70**	.12	-.16	-.46**	-.47**	.06	-.27*
13.	LSAT Intercept			.01	.12	-.19	.03	-.03	.04
14.	LGZV Slope				-.79**	.10	.15	.27*	-.31*
15.	LGZV Intercept					-.18	.09	-.14	.40**
16.	LGZO Slope						.70**	. .	. .
17.	LGZO Intercept							. .	. .
18.	LBRT Slope								-.63**
19.	LBRT Intercept								

** $p \leq .01$
* $p < .05$
[a]Sample sizes varied from 31 to 127.

angle and viewing distances on accuracy or mean latency of responses. An improved but more costly testing situation would have a display at each station with examinees allowed to proceed at individual rates.

For accuracy scores, split-half reliabilities ranged from .62 to .93, with an average of .76. For mean latency the range was .78 to .95, with an average of .90. Reliabilities of latency scores were typically higher than reliabilities of the corresponding accuracy scores. These reliabilities are acceptable, but could probably be improved by including additional reliable items. Questions about reliability not answered by these studies concern (i) the test-retest reliabilities of mean accuracy and latency, (ii) the reliabilities of the derived measures, and (iii) the relationship of reliabilities to testing conditions such as the deadline for answering problems.

The pattern of correlations in each experiment generally followed that found by Johnson (10) for a quite different set of tests. For the four new tests under the conditions studied the following rules can be induced from the data. Measures of accuracy on spatial tests correlate significantly (average correlation was .40). Measures of latency on spatial tests correlate significantly (average correlation was .53). For practical purposes, measures of accuracy do not correlate significantly with measures of latency (average correlation was -.20). The latter rule had several exceptions in Experiment I, but they did not replicate in Experiment II.

This pattern of correlation is consistent with two very different interpretations of the mean latency scores. One is that spatial processing has two distinct components, one measured by speed, the other by accuracy. Another interpretation is that latency reflects a general characteristic such as perceptual speed, motivation, or general intelligence that is not peculiar to spatial processing. Latency scores should be related to other types of variables in future research to determine if they reflect a general or a spatial process.

As indicated by accuracy scores, the content of a test appears to be preserved from a total-timed, paper-and-pencil, multiple-choice format to an item-timed, slide-projected, binary-choice format. That statement is supported by the correlations between accuracy on the standard and redesigned forms of the tests.

For extension of the theoretical work in spatial processing, the conclusions are limited by the methods employed in these studies. The data suggest that Spatial Orientation can be interpreted as a form of concept verification in which each of the three spatial dimensions of a figure is serially checked against the concept of what the figure ought to be. This descriptive model had more support from data on the LSAT than it did from data on the LGZO. For each test, approximately 25 percent of the examinees produced data inconsistent with the Spatial Orientation model.

Several explanations can be advanced for this discrepancy. Some examinees may process spatial orientation problems differently (i.e., parallel rather than serial matching) or possibly they are not at all systematic in their approach. On the other hand, the data of individual examinees may not be reliable enough to indicate whether they are behaving according to the Spatial Orientation model. Given the high error rates and correspondingly low number of correct "No" responses for each examinee in these experiments, the latter alternative deserves serious consideration. These questions have to be resolved using a different experimental procedure.

Spatial Visualization was found to have properties analogous to physically turning an object in space. It was found for the LGZV that a greater number of mental turns required a correspondingly greater amount of time to solve the problem. For the LBRT, the angular extent of a single mental rotation was systematically related to the time to solve the problem. In fact, the LGZV and LBRT were the most similar pair of tests, correlating significantly on every measure.

Additional measures of spatial processing were derived from these theoretical ideas. For each examinee on each test, a slope and an intercept was computed for the best-fitting line relating the latency of response to characteristics of the spatial problems. The highest correlations among the derived measures occur between slopes and their corresponding intercepts. This is disappointing because there is no theoretical reason to expect the dependency of these parameters. In the Spatial Visualization and Spatial Orientation models, the parameters characterize distinct processes, but in the data they have about 50 per cent of their variance in common. The relationship is the same in each case, low slope values being related to high intercepts.

This relationship may be artificially induced by using a time limit for test items. For example, an examinee with a long input time may have to make rapid mental rotations to answer the item before the time limit. Another possible explanation is that errors in estimates of slopes cause systematic errors in estimates of intercepts. With high error rates and large item differences, split-half reliability is impractical to use for slopes and intercepts. To obtain estimates of reliabilities, a test-retest paradigm ought to be employed. Using practiced subjects and a different method of estimating intercepts, Snyder (14) found a different pattern of slope intercept correlation for a test similar to the LBRT. The question of selecting derived measures rather than simpler measures can only be resolved by further research examining the reliability and validity of each measure.

CONCLUSIONS

1. Using suitably designed tests, the collection of latency measures of spatial ability in a group testing situation proved feasible.

2. Mean latencies obtained under the conditions studied had an average reliability of .90, while reliabilities of accuracy measures averaged .76. The reliability of accuracy measures was probably curtailed by the use of the two-choice procedure and relatively small numbers of items per test.

3. The following pattern of correlation was observed. (i) Accuracy scores correlated across all tests. Correlations near the level of alternate-form reliability were observed between accuracy scores on a standard test and its redesigned counterpart. (ii) Mean latency of correct responses correlated significantly across all tests. (iii) Correlations among latency and accuracy were negative and generally of low magnitude. Consequently, accuracy and latency scores give consistent but distinct information about examinees.

4. In the terms of Information Processing, Spatial Orientation appeared to be a form of concept verification in which examinees serially check the three spatial dimensions of a figure against their concept of what the figure should be. Spatial Visualization appeared to have properties analogous to physically turning an object in space, so that problems requiring a greater number of turns or turns of greater length required more time to solve.

5. On the basis of these models, derived measures of spatial processing were selected and analyzed. Conclusions concerning these measures are limited. Further work ought to rigorously test the spatial processing models, and establish the reliability and validity of derived measures.

REFERENCES

1. Ambler, R. K., and Smith, M. J. Differentiating Aptitude Factors Among Current Aviation Specialties. NAMRL Tech. Report No. 1207, Naval Aerospace Medical Research Laboratory, Pensacola, Florida, August, 1974.

2. Cooper, L. A., and Shepard, R. N. Chronometric Studies of the Rotation of Mental Images. In W. G. Chase (Ed.) Visual Information Processing. New York: Academic Press, 1973.

3. Egan, D. E., and Greeno, J. G. Theory of rule induction. In L. W. Gregg (Ed.) Knowledge and Cognition. Potomac, Maryland: Lawrence Earlbaum Associates, 1974.

4. French, J. W. The description of aptitude and achievement tests in terms of rotated factors. Psychometric Monographs, 1951, No. 5.

5. Gannon, F. B. Final Report for Contract Number Nonr-4688(00). Educational Testing Service, March, 1969.

6. Guilford, J. P. The Nature of Human Intelligence. New York: McGraw-Hill, 1969.

7. Guilford, J. P., and Lacey, J. I. (Eds.) Printed Classification Tests. Army Air Forces Aviation Psychology Program, Research Report No. 5, 1947.

8. Guilford, J. P., and Zimmerman, W. S. The Guilford-Zimmerman Aptitude Survey. Beverly Hills: Sheridan Supply Co., 1947.

9. Guilford, J. P., and Zimmerman, W. S. The Guilford-Zimmerman Aptitude Survey. Journal of Applied Psychology, 1948, 32, 24-34.

10. Johnson, L. C. Speed and confidence of judgment as psychological variables. Technical Report 4, Department of Psychology, Stanford University, 1954.

11. Kelly, T. L. Crossroads In The Mind of Man: a Study of Differentiable Mental Abilities. Stanford, Calif.: Stanford, 1928.

12. Shepard, R. N., and Feng, C. A chronometric study of mental paper folding. Cognitive Psychology, 1972, 3, 228-243.

13. Shepard, R. N., and Metzler, J. Mental rotation of three-dimensional objects. Science, 1971, 171, 701-703.

14. Snyder, C. R. Individual differences in imagery and thought. Unpublished doctoral dissertation, University of Oregon, June, 1972.

15. Thurstone, L. L. Primary mental abilities. Psychometric Monographs, 1938, No. 1.

Figure Captions

Fig. 1 An item from the Spatial Apperception Test (top), and an item from the redesigned test (bottom).

Fig. 2 An item from the Guilford-Zimmerman Spatial Visualization Test (top) and an item from the redesigned test (bottom).

Fig. 3 An item from the Guilford-Zimmerman Spatial Orientation Test (top), and an item from the redesigned test (bottom).

Fig. 4 A "YES" item (top) and a "NO" item (bottom) from the Block Rotation test.

Fig. 5 Information processing model of Spatial Visualization task.

Fig. 6 Mean latency for correct and wrong answers to items requiring <u>n</u> turns on the LGZV (Experiment I).

Fig. 7 Mean latency for correct "YES" responses as a function of required angle of rotation to achieve a match on the LBRT.

Fig. 8 Information processing model of Spatial Orientation task.

Fig. 9 Mean latency for correct "NO" responses on the LSAT and LGZO as a function of the number of discrepancies between the correct spatial concept and the figure presented (Experiment I).

Egan, D. E., LT, MSC, USNR (Ph.D) February 1976

ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING. NAMRL-
Pensacola, Fla.: Naval Aerospace Medical Research Laboratory. February 1976

The processing of visually presented spatial information is a critical component of the activities performed by pilots and aircrewmen. In particular, Radar Intercept Officers and Air Control Officers must make rapid and accurate spatial judgments. It is likely that variation in the ability to process spatial information accounts for some of the undesirable variations in the performance of these jobs. Previous research using conventional or "accuracy" scoring for paper-and-pencil tests has identified two "spatial factors" (Spatial Orientation and Spatial Visualization) that are valid predictors of success in pilot and navigator training programs. Recent experimental work has used the latency of response to spatial problems to analyze the mental processing of spatial information. The present studies combine these approaches by investigating both accuracy and latency scores as measures of the ability to process spatial information. Spatial test items were redesigned to be suitable for collecting latency as well as accuracy scores. In two experiments four new spatial tests were administered to groups of U. S. Navy pilot and Flight Officer Candidates. The psychometric properties of latency and accuracy scores from those tests were determined. Informal tests of several hypotheses about spatial processing were carried out. Derived measures of spatial processing were proposed and analyzed. Response latency scores are both feasible and desirable for assessing the ability to process spatial information. Latency scores were highly reliable and correlated across different spatial tests. Accuracy scores were somewhat less reliable, but correlated predictably across tests. Interestingly, latency and accuracy were virtually independent measures. Tentative support was found for a model of Spatial Orientation patterned after theories of concept verification. Spatial Visualization appeared to be a continuous process similar to physically turning an object in space. Measures of spatial processing based on those models correlated in a consistent pattern.

Spatial ability.
Response latency.
Response speed.
Computerized testing.

Egan, D. E., LT, MSC, USNR (Ph.D) February 1976

ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING. NAMRL-
Pensacola, Fla.: Naval Aerospace Medical Research Laboratory. February 1976

The processing of visually presented spatial information is a critical component of the activities performed by pilots and aircrewmen. In particular, Radar Intercept Officers and Air Control Officers must make rapid and accurate spatial judgments. It is likely that variation in the ability to process spatial information accounts for some of the undesirable variations in the performance of these jobs. Previous research using conventional or "accuracy" scoring for paper-and-pencil tests has identified two "spatial factors" (Spatial Orientation and Spatial Visualization) that are valid predictors of success in pilot and navigator training programs. Recent experimental work has used the latency of response to spatial problems to analyze the mental processing of spatial information. The present studies combine these approaches by investigating both accuracy and latency scores as measures of the ability to process spatial information. Spatial test items were redesigned to be suitable for collecting latency as well as accuracy scores. In two experiments four new spatial tests were administered to groups of U. S. Navy pilot and Flight Officer Candidates. The psychometric properties of latency and accuracy scores from those tests were determined. Informal tests of several hypotheses about spatial processing were carried out. Derived measures of spatial processing were proposed and analyzed. Response latency scores are both feasible and desirable for assessing the ability to process spatial information. Latency scores were highly reliable and correlated across different spatial tests. Accuracy scores were somewhat less reliable, but correlated predictably across tests. Interestingly, latency and accuracy were virtually independent measures. Tentative support was found for a model of Spatial Orientation patterned after theories of concept verification. Spatial Visualization appeared to be a continuous process similar to physically turning an object in space. Measures of spatial processing based on those models correlated in a consistent pattern.

Spatial ability.
Response latency.
Response speed.
Computerized testing.

Egan, D. E., LT, MSC, USNR (Ph.D) February 1976

ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING. NAMRL-
Pensacola, Fla.: Naval Aerospace Medical Research Laboratory. February 1976

The processing of visually presented spatial information is a critical component of the activities performed by pilots and aircrewmen. In particular, Radar Intercept Officers and Air Control Officers must make rapid and accurate spatial judgments. It is likely that variation in the ability to process spatial information accounts for some of the undesirable variations in the performance of these jobs. Previous research using conventional or "accuracy" scoring for paper-and-pencil tests has identified two "spatial factors" (Spatial Orientation and Spatial Visualization) that are valid predictors of success in pilot and navigator training programs. Recent experimental work has used the latency of response to spatial problems to analyze the mental processing of spatial information. The present studies combine these approaches by investigating both accuracy and latency scores as measures of the ability to process spatial information. Spatial test items were redesigned to be suitable for collecting latency as well as accuracy scores. In two experiments four new spatial tests were administered to groups of U. S. Navy pilot and Flight Officer Candidates. The psychometric properties of latency and accuracy scores from those tests were determined. Informal tests of several hypotheses about spatial processing were carried out. Derived measures of spatial processing were proposed and analyzed. Response latency scores are both feasible and desirable for assessingthe ability process spatial information. Latency scores were highly reliable and correlated across different spatial tests. Accuracy scores were somewhat less reliable, but correlated predictably across tests. Interestingly, latency and accuracy were virtually independent measures. Tentative support was found for a model of Spatial Orientation patterned after theories of concept verification. Spatial Visualization appeared to be a continuous process similar to physically turning an object in space. Measures of spatial processing based on those models correlated in a consistent pattern.

Spatial ability.
Response latency.
Response speed.
Computerized testing.

Egan, D. E., LT, MSC, USNR (Ph.D) February 1976

ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING. NAMRL-
Pensacola, Fla.: Naval Aerospace Medical Research Laboratory. February 1976

The processing of visually presented spatial information is a critical component of the activities performed by pilots and aircrewmen. In particular, Radar Intercept Officers and Air Control Officers must make rapid and accurate spatial judgments. It is likely that variation in the ability to process spatial information accounts for some of the undesirable variations in the performance of these jobs. Previous research using conventional or "accuracy" scoring for paper-and-pencil tests has identified two "spatial factors" (Spatial Orientation and Spatial Visualization) that are valid predictors of success in pilot and navigator training programs. Recent experimental work has used the latency of response to spatial problems to analyze the mental processing of spatial information. The present studies combine these approaches by investigating both accuracy and latency scores as measures of the ability to process spatial information. Spatial test items were redesigned to be suitable for collecting latency as well as accuracy scores. In two experiments four new spatial tests were administered to groups of U. S. Navy pilot and Flight Officer Candidates. The psychometric properties of latency and accuracy scores from those tests were determined. Informal tests of several hypotheses about spatial processing were carried out. Derived measures of spatial processing were proposed and analyzed. Response latency scores are both feasible and desirable for assessingthe ability process spatial information. Latency scores were highly reliable and correlated across different spatial tests. Accuracy scores were somewhat less reliable, but correlated predictably across tests. Interestingly, latency and accuracy were virtually independent measures. Tentative support was found for a model of Spatial Orientation patterned after theories of concept verification. Spatial Visualization appeared to be a continuous process similar to physically turning an object in space. Measures of spatial processing based on those models correlated in a consistent pattern.

Spatial ability.
Response latency.
Response speed.
Computerized testing.

Egan, D. E., LT, MSC, USNR (Ph.D) February 1976

ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING. NAMRL
Pensacola, Fla.: Naval Aerospace Medical Research Laboratory. February 1976

The processing of visually presented spatial information is a critical component of the activities performed by pilots and aircrewmen. In particular, Radar Intercept Officers and Air Control Officers must make rapid and accurate spatial judgments. It is likely that variation in the ability to process spatial information accounts for some of the undesirable variations in the performance of these jobs. Previous research using conventional or "accuracy" scoring for paper-and-pencil tests has identified two "spatial factors" (Spatial Orientation and Spatial Visualization) that are valid predictors of success in pilot and navigator training programs. Recent experimental work has used the latency of response to spatial problems to analyze the mental processing of spatial information. The present studies combine these approaches by investigating both accuracy and latency scores as measures of the ability to process spatial information. Spatial test items were redesigned to be suitable for collecting latency as well as accuracy scores. In two experiments four new spatial tests were administered to groups of U. S. Navy pilot and Flight Officer Candidates. The psychometric properties of latency and accuracy scores from those tests were determined. Informal tests of several hypotheses about spatial processing were carried out. Derived measures of spatial processing were proposed and analyzed. Response latency scores are both feasible and desirable for assessing the ability to process spatial information. Latency scores were highly reliable and correlated across different spatial tests. Accuracy scores were somewhat less reliable, but correlated predictably across tests. Interestingly, latency and accuracy were virtually independent measures. Tentative support was found for a model of Spatial Orientation patterned after theories of concept verification. Spatial Visualization appeared to be a continuous process similar to physically turning an object in space. Measures of spatial processing based on those models correlated in a consistent pattern.

Spatial ability.
Response latency.
Response speed.
Computerized testing.

Egan, D. E., LT, MSC, USNR (Ph.D) February 1976

ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING. NAMRL-
Pensacola, Fla.: Naval Aerospace Medical Research Laboratory. February 1976

The processing of visually presented spatial information is a critical component of the activities performed by pilots and aircrewmen. In particular, Radar Intercept Officers and Air Control Officers must make rapid and accurate spatial judgments. It is likely that variation in the ability to process spatial information accounts for some of the undesirable variations in the performance of these jobs. Previous research using conventional or "accuracy" scoring for paper-and-pencil tests has identified two "spatial factors" (Spatial Orientation and Spatial Visualization) that are valid predictors of success in pilot and navigator training programs. Recent experimental work has used the latency of response to spatial problems to analyze the mental processing of spatial information. The present studies combine these approaches by investigating both accuracy and latency scores as measures of the ability to process spatial information. Spatial test items were redesigned to be suitable for collecting latency as well as accuracy scores. In two experiments four new spatial tests were administered to groups of U. S. Navy pilot and Flight Officer Candidates. The psychometric properties of latency and accuracy scores from those tests were determined. Informal tests of several hypotheses about spatial processing were carried out. Derived measures of spatial processing were proposed and analyzed. Response latency scores are both feasible and desirable for assessing the ability to process spatial information. Latency scores were highly reliable and correlated across different spatial tests. Accuracy scores were somewhat less reliable, but correlated predictably across tests. Interestingly, latency and accuracy were virtually independent measures. Tentative support was found for a model of Spatial Orientation patterned after theories of concept verification. Spatial Visualization appeared to be a continuous process similar to physically turning an object in space. Measures of spatial processing based on those models correlated in a consistent pattern.

Spatial ability.
Response latency.
Response speed.
Computerized testing.

Egan, D. E., LT, MSC, USNR (Ph.D) February 1976

ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING. NAMRL
Pensacola, Fla.: Naval Aerospace Medical Research Laboratory. February 1976

The processing of visually presented spatial information is a critical component of the activities performed by pilots and aircrewmen. In particular, Radar Intercept Officers and Air Control Officers must make rapid and accurate spatial judgments. It is likely that variation in the ability to process spatial information accounts for some of the undesirable variations in the performance of these jobs. Previous research using conventional or "accuracy" scoring for paper-and-pencil tests has identified two "spatial factors" (Spatial Orientation and Spatial Visualization) that are valid predictors of success in pilot and navigator training programs. Recent experimental work has used the latency of response to spatial problems to analyze the mental processing of spatial information. The present studies combine these approaches by investigating both accuracy and latency scores as measures of the ability to process spatial information. Spatial test items were redesigned to be suitable for collecting latency as well as accuracy scores. In two experiments four new spatial tests were administered to groups of U. S. Navy pilot and Flight Officer Candidates. The psychometric properties of latency and accuracy scores from those tests were determined. Informal tests of several hypotheses about spatial processing were carried out. Derived measures of spatial processing were proposed and analyzed. Response latency scores are both feasible and desirable for assessingthe ability process spatial information. Latency scores were highly reliable and correlated across different spatial tests. Accuracy scores were somewhat less reliable, but correlated predictably across tests. Interestingly, latency and accuracy were virtually independent measures. Tentative support was found for a model of Spatial Orientation patterned after theories of concept verification. Spatial Visualization appeared to be a continuous process similar to physically turning an object in space. Measures of spatial processing based on those models correlated in a consistent pattern.

Spatial ability.
Response latency.
Response speed.
Computerized testing.

Egan, D. E., LT, MSC, USNR (Ph.D) February 1976

ACCURACY AND LATENCY SCORES AS MEASURES OF SPATIAL INFORMATION PROCESSING. NAMRL-
Pensacola, Fla.: Naval Aerospace Medical Research Laboratory. February 1976

The processing of visually presented spatial information is a critical component of the activities performed by pilots and aircrewmen. In particular, Radar Intercept Officers and Air Control Officers must make rapid and accurate spatial judgments. It is likely that variation in the ability to process spatial information accounts for some of the undesirable variations in the performance of these jobs. Previous research using conventional or "accuracy" scoring for paper-and-pencil tests has identified two "spatial factors" (Spatial Orientation and Spatial Visualization) that are valid predictors of success in pilot and navigator training programs. Recent experimental work has used the latency of response to spatial problems to analyze the mental processing of spatial information. The present studies combine these approaches by investigating both accuracy and latency scores as measures of the ability to process spatial information. Spatial test items were redesigned to be suitable for collecting latency as well as accuracy scores. In two experiments four new spatial tests were administered to groups of U. S. Navy pilot and Flight Officer Candidates. The psychometric properties of latency and accuracy scores from those tests were determined. Informal tests of several hypotheses about spatial processing were carried out. Derived measures of spatial processing were proposed and analyzed. Response latency scores are both feasible and desirable for assessingthe ability process spatial information. Latency scores were highly reliable and correlated across different spatial tests. Accuracy scores were somewhat less reliable, but correlated predictably across tests. Interestingly, latency and accuracy were virtually independent measures. Tentative support was found for a model of Spatial Orientation patterned after theories of concept verification. Spatial Visualization appeared to be a continuous process similar to physically turning an object in space. Measures of spatial processing based on those models correlated in a consistent pattern.

Spatial ability.
Response latency.
Response speed.
Computerized testing.