An Investigation of Two Hypotheses Regarding The Nature of the Spatial-Rlations and Visualization Factors

William B. Michael, Wayne S. Zimmerman and J.P. Guilford

Educational and Psychological Measurement 1950 10: 187
DOI: 10.1177/001316445001000202

AN INVESTIGATION OF TWO HYPOTHESES REGARDING THE NATURE OF THE SPATIAL-RELATIONS AND VISUALIZATION FACTORS[1]

WILLIAM B. MICHAEL
Princeton University
and
WAYNE S. ZIMMERMAN and J. P. GUILFORD
University of Southern California

PRIMARILY as a consequence of the factorial analyses of tests of intellectual abilities, the construct of a spatial and/or visual ability amenable to psychological measurement has received increasing attention in recent years. During the past twenty-one years, at least a score of investigators have identified in their writings a space factor. In a pioneer study, Thurstone (16) included among his seven primary mental abilities a factor labelled S, which he characterized as a "facility in spatial and visual imagery,"—a factor which he likened to the spatial or visual group factor found by Kelley (13) in earlier experiments. The same factor was identified in other studies carried out subsequently by Thurstone (17) and by Thurstone and Thurstone (18).

During World War II members of the psychological research units of the Army Air Forces devoted a considerable amount of time and effort to the development of tests of the "spatial-visual" type to be used in the selection of men for air-crew positions. Several factorial studies which have been described in a research report of the AAF Aviation Psychology Research Program, edited by Guilford (3), have indicated that the variance associated with Thurstone's spatial-visualization factor may be separated into two apparently independent factors identified to be spatial relations and visualization (visual manipulation). In fact, in addition to these two factors (abbreviated by the symbols S_1 and Vz), two other less definite space factors, S_2 and S_3, and a factor tentatively identified as visual memory, also appeared in several analyses.

In two recent studies both Fruchter (2) and Dudek (1) have found separate factors of spatial relations and visualization. In his investigation as to the nature of verbal fluency Fruchter reanalyzed a sub-matrix of twenty tests selected from the battery of fifty-seven variables employed by Thurstone in his classical study previously cited (16). He found two independent factors which he described as being spatial-relations and visualization.

Referring to the same Thurstone study, Zimmerman (3) pointed out that further rotations of the residual axis (Number XII) with other axes which defined meaningful factors would produce a promising factor of visualization. Just recently, Zimmerman (in his unpublished doctoral dissertation) has rerotated the twelve centroid axes for all fifty-seven variables and has confirmed his initial belief that both a spatial-relations and a visualization factor would appear.

Problem

THE purpose of the investigation was to test the validity of two (apparently unrelated) hypotheses that purport to represent differences in the psychological properties of the factors of spatial-relations and visualization as reflected by corresponding differences, both in the respective contents of two types of tasks and in the respective work procedures required of the subjects for successful completion of them. Each type of task consisted of a group of three tests. Within each of the two groups of tests employed in the study there appeared to be not only a similarity in the format of the test items, but also a common approach or operation demanded of the examinee.

In broad outline the plan followed in the investigation was to incorporate within a test battery two groups of tests which the investigators believed to be representative of the psycho-

[1] The first-mentioned author wishes to express his sincere appreciation to the Social Science Research Council for kindly making available a grant-in-aid for the completion of this investigation. The authors are indebted to Professor L. L. Thurstone, who generously granted permission to have several of his tests offprinted in order that they might be included within the battery. Grateful acknowledgment is made to the staff of the Department of Psychology at Rutgers University for their interest in the study, their cooperation in making subjects available, and their assistance in administering a number of the tests. To all Rutgers students who participated, special thanks are extended.

logical operations involved in the hypothetical statements as to the nature of the spatial-relations factor and of the visualization factor. In the selection of each test to be incorporated within a group, introspection was freely employed as an aid to the determination of the psychological processes used in the subjects' performance upon a test—the same processes supposedly as those indicated in the relevant hypotheses. To these six tests (i.e., two groups, each consisting of three tests) were added eight reference tests of fairly well-known factorial content to aid in the identification of those portions of variance in the six tests that were associated with other factors such as verbality, numerical facility, reasoning, and perceptual speed. The inclusion of other factor tests served not only to identify what probably without their presence would be large amounts of specific variance within each of the six tests, but also to indicate the relative degree of purity of each of these six tests with respect to the function it was hypothesized to measure.[2]

A sufficient, though not necessary, condition for the tenability of each of the hypotheses, would be that in the factor-analysis procedure each of the two groups of three tests would define a factor. Moreover, this factor should not appear to be weighted in other tests of the battery that were selected to measure other factors. If one or more tests within either group should be weighted substantially in variance associated with another factor, the evidence for the corresponding hypothesis would be less clear-cut, but not necessarily lacking. It would be quite possible, if not almost certain, that one or more of the three tests within a given group might be factorially complex. At the same time, however, all three tests within a given group might contain substantial amounts of variance in one factor that did not appear in any of the other eleven tests.

Hypotheses

The factor of spatial relations was hypothesized to represent the ability to comprehend the *arrangement* of elements within

[2] It was also thought to be very desirable to determine whether tests of the type used by the AAF and Thurstone's tests held in common factors identified as being the same. This is the first study the writers know of that will serve to check upon the belief that many of the Thurstone primary abilities and the AAF factors are identical. Only the Thurstone space factor is here called into question.

a visual stimulus pattern, primarily with reference to the human body. Thus, an important implication in the ability to perceive spatial arrangements is that the subject is able to distinguish whether one object is higher or lower, left or right, or nearer or farther than another within the same field. Through the presentation of two simulated views of a stimulus pattern, a test item may be constructed such that there is a systematic relationship between the order of elements within the first spatial pattern (the stimulus component of a test item) and the order of elements within the second pattern (the response component of a test item).

For example, in Thurstone's *Cubes* test the examinee is asked to recognize whether the designs on the sides of a second cube can hold the same relationship to one another as they do on the first cube. By noticing within each cube the left-right, top-bottom, and front-back interrelationships of the faces, the subject is able in each item to refer the locations of three designs on three exposed faces of one cube to the locations of designs on the faces of the other cube. In Thurstone's *Flags* test the examinee is required to tell whether the exposed faces of two American flags of identical size can represent the same side of the flag. Relating corresponding left-right and top-bottom boundaries (outlines) of the two flags appears to be an important aspect of the solution. Similarly, in Guilford and Zimmerman's test of *Spatial Orientation* a premium is placed upon the examinee's maintaining the correct relationship of objects to one another in background scenery that has been viewed twice from a motorboat—first before and then after its prow has moved up or down and/or left or right. In the test the examinee is asked to determine the relative amount and direction of movement of the boat corresponding to changes in the two views of the background setting.

The factor of visualization was hypothesized to represent an ability that requires the mental *manipulation* of visual images. In contrast to another factor identified as visual memory (3), which appears to be a static or reproductive form of visualization, the factor referred to as visual manipulation, or simply visualization, is dynamic. This visual manipulative ability appears to be present in the solution of problems in which the in-

dividual finds it necessary mentally to move, rotate, turn, twist, or invert one or more objects. Following the performance of the presented manipulation the individual is required to recognize the new position, location, or changed appearance of the object or objects.

Three tests selected to yield evidence for the second hypothesis included two by Thurstone, *Punched Holes* and *Form Board*, and one by Guilford and Zimmerman, *Spatial Visualization*. In the test of *Punched Holes* the examinee is presented a symbolic representation of a folded sheet of paper into which one or more holes have been punched and is required to imagine where the holes will be when the sheet is unfolded. In the second Thurstone test the examinee apparently finds it necessary in each item mentally to turn, rotate, or invert two or more flat geometric figures in such a way that they can be placed together to fit within the outline of a larger geometric figure. In each of the tests, the examinee is asked to record the final positions respectively of the holes and of the geometric figures. In the test of *Spatial Visualization* the subject is required mentally to turn, tilt, or rotate a three-dimensional object—an alarm clock—drawn on a sheet of paper into a final position according to written instructions. As alternative responses the pictures of the clock are presented in five positions, one of which is correct. (A more detailed description of these three tests follows in the next section.)

Whereas in the two Thurstone tests the examinee is required to draw in his solution to the problem, in the third test he merely selects as his solution one of five choices presented. It is quite likely that in addition to measuring visual manipulative ability other factors are involved in the three tests—factors reflecting the manner in which responses to the items are recorded.

Another important difference in the nature of the psychological processes hypothesized for the spatial relations and visualization factors was that of speed of response. As indicated by findings in the AAF Aviation Psychology Program, the tests thought to measure the spatial relations factor were administered with fairly short time limits, but those tests thought to measure visualization were given with fairly liberal time allowances. The spatial relations factor was considered to demand a fairly rapid decision on the part of the examinee as to the spatial position of objects with reference to his own location; whereas, the visualization factor was believed to be represented in problems requiring a more deliberate and less automatic approach. In part, such a distinction may be a function of the complexity of a task (i.e., the number of steps entering into the performance of an item), the more complex tasks requiring visualization for their solution.

Concerning the psychological properties of spatial-relations and visualization factors, one other important difference has been suggested in the work of one of the psychological research units of the AAF, as follows:

> The idea for Flight Orientation [a test] was proposed at the time Aerial Orientation [another test] was being developed. It was hypothesized (1) that the ability visually to maneuver an airplane as if from a position outside the cockpit is a manipulatory-visualization ability and (2) that the ability to imagine maneuvers taking place as if the examinee were within the cockpit is a spatial-orientation ability.
>
> The Aerial Orientation test utilized cockpit views of outside terrain to be matched with depicted plane attitudes; the visualization-of-maneuvers tests involved only views of airplanes seen from a position outside of the cockpit. . . . Flight Orientation was designed to fulfill the requirements of the indicated variation—a test that would utilize only cockpit views of outside terrain. From hypotheses given above, it follows that Aerial Orientation should measure a combination of manipulatory-visualization and spatial-orientation abilities, while Flight Orientation should be a purer measure of the ability to orient in space (3).

That the two groups of tests selected for investigating the validity of the hypotheses may actually contain variance in both the spatial-relations and visualization factors would not be surprising, inasmuch as many subjects on the basis of their own introspective reports revealed that they made use of the two psychological processes associated with the respective hypotheses in tests selected to represent the implications of only one hypothesis. For example, if in the *Flags* test the subject is able, so to speak, to pick up the flag, move it, turn it about as if he actually has a model in his hands, then visualization is believed to be dominant. On the other hand, if the subject is

concerned primarily with the left-right and top-bottom orientation of edges of flags with respect to his own position, or if he has to move himself to a different position as in cocking his head to one side, then a spatial factor is believed to be more prominent.

Similarly, in the *Cubes* test if the subject reports he picks up the first cube and rotates it into a final position which matches (or cannot match) the second cube, then the visualization process is dominant. However, if he attempts primarily to interrelate the positions of the sides of the cubes with respect to his own position, or if he appears to project himself amidst the cubes as if he were walking about them and relating the locations of various sides with respect to his own position, then the spatial-relations factor is probably operative. It may well be that in the spatial-relations factor empathy plays an important role in the relating of the position of objects to one's own location; whereas in visualization the individual obtains first from a distance an overall view of the objects to be manipulated and then employs perhaps some rather restricted kinesthetic imagery in the imagined use of hands for moving the objects into their required positions.

Despite the apparent differences in approach employed by many subjects, it did appear that the two groups of tests chosen represented reasonably well a distinction between the psychological processes hypothesized. If a test did involve to a substantial degree the use of two or more psychological abilities, it was thought that the factor-analysis procedure would reveal such a fact.

Tests

In Table 1 are presented the names of the fourteen pencil-and-paper tests employed in the battery, the maximum number of items that could be attempted, the plan followed with respect to "speed" or "power" time-limit, the actual working time allowed, and the scoring formula used. The numbering of the tests in the tables, as well as in the following description of content and procedure, corresponds to the order of administration. During the first, second, third, and fourth testing periods, respectively, the following groups of tests were ad-

ministered: 1, 2, 3, and 4; 5, 6, 7, 8, and 9; 10 and 11; 12, 13, and 14. An ample number of practice exercises preceded the main body of each test. Further information concerning several of the tests may be found both in a manual (11) and in the literature (12, 16, 18, 20). It is believed, however, that the descriptions given will suffice for the interpretation of the factors to be presented.

1. *Guilford-Zimmerman Verbal Comprehension.*—This is a vocabulary test in which the examinee is required in each item to choose among five words, all matched with respect to difficulty, the one word which most closely approximates the meaning of the stimulus word. Items increase in difficulty progressively from the beginning to the end. Even numbered items were omitted. Responses were recorded on a separate answer sheet. Most examinees attempted all items.

2. *Guilford-Zimmerman General Reasoning.*—This test is composed of arithmetical-reasoning problems similar to those encountered in courses in general mathematics, elementary algebra, and intermediate algebra. Diagrams accompany a few of

TABLE 1

The Test Battery: Descriptive Data

Name of Test	Number of Items	Timing Plan (Speed or Power)	Working Time	Scoring Formula
1. Guilford-Zimmerman Verbal Comprehension	36	Power	10 min.	R-W/4
2. Guilford-Zimmerman General Reasoning	14	Power	13 min.	R-W/4
3. Guilford-Zimmerman Numerical Operations	120	Speed	5 min.	R-W
4. Guilford-Zimmerman Perceptual Speed	48	Speed	3 min., 45 sec.	R-W
5. Guilford-Zimmerman Spatial Orientation	60	Speed	8 min.	R-W/4
6. Thurstone [Verbal] Completion	30	Power	7 min.	R-W/4
7. Thurstone Number Series	20	Power	8 min.	R
8. Thurstone Identical Forms	40	Speed	3 min., 15 sec.	R-W
9. Thurstone Cubes	42	Speed	5 min.	R-W
10. Thurstone Flags	48	Speed	4 min.	R-W
11. Guilford-Zimmerman Spatial Visualization	40	Power (limited)	15 min.	R-W/4
12. Thurstone Punched Holes	10	Power	7 min.	R
13. Thurstone Pattern Analogies	20	Power	10 min.	R-W/4
14. Thurstone Form Board	28	Power	7 min.	R

the problems. Numerical work is kept to a minimum. Five multiple-choice responses are presented with each problem statement. Items increase in difficulty level progressively from the beginning to the end. Even-numbered items were omitted. Responses were recorded on a separate answer sheet. Most examinees attempted all assigned items.

3. *Guilford-Zimmerman Numerical Operations.*—This test is in four parts, consisting of numerous simple problems (of about the same difficulty level) involving respectively the four fundamental operations of addition, subtraction, multiplication and division. Emphasis is placed in the directions upon the need for both accuracy and speed of work. Subjects were told to begin with the part upon addition, to work every item, and to go as far as possible in the allotted time. Only a few subjects reached the fourth section upon division. Responses to the items were printed in spaces on the test booklet adjacent to the problems.

4. *Guilford-Zimmerman Perceptual Speed.*—This test requires the examinee to match a visual object of a familiar shape and of detailed design with one of five other visual objects of a common category (e.g., automobiles, boats, hats, shoes). Four of the five response objects resemble rather closely the stimulus object, but differ from it in certain minor details of shape and/or design. For each common category two parallel sets of visual objects—four stimulus and five response objects—are arranged in two parallel columns. To each one of the four stimulus figures in the first column corresponds one of the five response figures. Thus, four responses are scored for each item of homogeneous content. All items represent a low level of difficulty. Answers to the items were marked on the test booklet in spaces adjacent to each stimulus object. The examinees were told to go as far as possible in the allotted time. No examinee finished.

5. *Guilford-Zimmerman Spatial Orientation.*—This test requires an examinee to determine how the position of a boat has changed in a second picture from its initial position in a first picture. In each picture the prow of the motorboat, in which the examinee is told to pretend to be riding, is shown along with background scenery consisting of water, or a silhouetted shore line, and in some instances of other boats intervening between the shore line and the prow of the motorboat, which is in the extreme foreground of the picture. In the sample problems described in detail in the directions, the position of the prow in the second picture, with respect to the spot of background sighted over it in the first picture, is taken as the primary reference guide for determination of the direction and amount of subsequent up-down and/or left-right motion of the boat. Movement is also indicated by accompanying shifts in the location of elements within the pattern of visible background scenery. The boat is actually stationary with respect to any forward-backward motion. To each set of two pictures five alternative responses are presented. Each response is represented by (1) a dot designating the aiming point, the initial spot in the background sighted right over the point of the prow in the first picture, and (2) an arc (of about 45°) representing the location of the prow in the second picture with reference to the aiming point. One of the five responses shows the correct change in position of the prow of the boat with respect to the aiming point. All examinees were instructed in the limited time allowed to attempt as many items as possible. As in all other speed tests, answers were recorded in the test booklet. The difficulty of the items tends to increase for items further removed from the beginning of the test. No one attempted every item.

6. *Thurstone* [Verbal] *Completion.*—This test is one adapted from the *Psychological Examination of the American Council on Education.* Representing, probably, a combination of verbal comprehension and verbal fluency, it presents for each item the definition of a word, the number of letters in the word, and five alternative letters (responses), one of which represents the initial letter of the defined word. Although the items differ considerably with respect to difficulty, most of the defined words are familiar to college students. Responses were recorded on the page of test items. Nearly every subject attempted all items.

7. *Thurstone Number Series.*—Found to be loaded in a factor identified by Thurstone as induction, this test requires the subject to determine a rule for each item. Numbers are presented in a row with two blanks inserted. The task is to find the mathematical principle by which the number series is formed and to insert in the blank that number which is appropriate. The difficulty level of items increases in relation to the position of the item from the beginning of the test. Responses were recorded on the test sheets in the blanks inserted at various positions within the different number series. Most of the subjects attempted all items. One point of credit was given to each blank correctly filled (two points per item being maximum score).

8. *Thurstone Identical Forms.*—This test resembles rather closely the fourth test, *Perceptual Speed*, in that the examinee selects from a row of five similar appearing figures that one which is exactly the same as the stimulus figure. Slight differences in color design and in shape appear among the five response figures. In this test the items are also homogeneous with respect to difficulty. The number corresponding to the sequential position of the response selected was recorded on the test page in a box to the right of the row of response objects. Only a few examinees reached the last few items.

9. *Thurstone Cubes.*—In this difficult test the subject is asked whether two drawings can represent the same cube on each face of which there is supposed to be a different design. In each of the two drawings the designs of three faces of the cubes are always exposed. If the two drawings can represent the same cube, a plus sign is placed in a blank square to the right of the two drawn cubes. If, on the other hand, the second drawing cannot represent the cube of the first drawing, then a negative sign is placed in the adjacent square. In the short time allowed no one attempted all items.

10. *Thurstone Flags.*—On this test two flag pictures, of the same size and of identical design, are presented occasionally in the same position, but generally in different positions. If the two drawings represent the same face of the flag, a plus sign is placed in a square on the test sheet just to the right of the two flags. If the two drawings represent opposite faces of the same flag, a minus sign is placed in the adjacent square. As in the test of *Cubes* the items were homogeneous with respect to difficulty. However, they were easy for most subjects. A few of the subjects attempted all items during the short period of time allowed.

11. *Guilford-Zimmerman Spatial Visualization.*—This is a test in which the examinee attempts to imagine the movement of a clock in space from an initial position to a final position as directed by a verbal statement. The test is divided into three parts. In the first part, one movement of the clock is required to effect the final position; in the second part, two movements are called for, and, in the third part, three movements are indicated by the directions accompanying each item. Three types of movements are required. Each type of movement refers to the revolution of the clock about an axis in one of three dimensions. The actual movement involves a revolution of the clock to the right or to the left a specified number of degrees. The word "turn" is used to designate a revolution about the base or the "6–12" axis where the numbers refer to the numerals representing hours on the clock. When the clock is tilted such that top moves either forward or backward, or in other words, when the clock is revolved about the "3–9" axis, the word "tilt" is employed. When the clock revolves about an axis perpendicular to its face, the word "rotate" is used. In the second part, two different types of movement are required, and six permutations of sequence of movement are used. In the third part, the same sequence of movements is followed in all items (rotate, tilt, and turn). Nearly all of the subjects failed to complete the entire test, but about 80 per cent attempted all items in the first two parts. Items were scored up to the point at which 67 per cent of the group attempted them.

12. *Thurstone Punched Holes.*—Each item in this test consists of a series of figures representing a square sheet of paper that has been folded by steps (as indicated by dotted lines) into smaller squares, rectangular, or triangular sizes. One or more holes are punched into the final folded form. The task for the subject is to imagine where the holes will be when the sheet is unfolded. As an aid to the subject's performance in the more difficult items one or more figures representing the appearance of the sheet of paper at intermediate stages of unfolding are presented. On the unfolded (square) sheet the subject indicates by drawing small circles where the holes will be. In the scoring of the item all holes must be properly spaced in relation to one another if credit is to be given. An item was scored right or

wrong (no partial credits were given). Nearly every subject completed all the items.

13. *Thurstone Pattern Analogies.*—Adapted from similar tests in the American Council on Education series, this test is composed of items each of which consists of eight figures. The first three (stimulus) figures are labelled A, B, C, and the next five (response) figures are designated *1*, *2*, *3*, *4*, and *5*. After the examinee determines the rule by which figure A is changed to figure B, he applies the rule to figure C and picks out among the five arabic numbered responses that one which satisfies the requirements of the problem. In the more complex items the examinee may frequently change his hypothesis as to the principle connecting A and B in view of limitations imposed by the nature of the five responses figures. In the time allowed, most subjects completed all items.

14. *Thurstone Form Board.*—Almost identical with the *Minnesota Form Board Test*, except for the inclusion of printed instructions and a practice exercise, this test consists of items made up of several two-dimensional pieces (colored black) of various geometrical shapes which the examinee attempts to fit together in an appropriate arrangement within a larger geometric form (uncolored figure within an outline). The subject draws lines within the large white (uncolored) design to show how the black pieces can be placed in order to fit within the outline. Extreme accuracy in drawing was not required, but the solution had to be indicated clearly. No partial credits were given. Although the items became increasingly difficult as one approached the end of the test, very few subjects failed to attempt all the items in the time allowed.

The Sample

To a group of 500 male students enrolled in a two-semester course in beginning psychology at Rutgers University the battery of fourteen pencil-and-paper tests was administered. Since four class periods, spread over the last part of the first semester and the first part of the second semester of the academic year, were required for completion of the project, many of the subjects were not present at all class sessions. Makeups were given in several instances. Complete results were obtained for 360

subjects. These individuals appeared to be a representative sample of the University student body in light of biographical information obtained from each student. Consisting of 220 freshmen and sophomores and 140 juniors and seniors majoring in virtually every department of the University, the sample was deemed satisfactory. Approximately 54 per cent of the subjects were veterans of World War II. The ages of the subjects ranged from 16 to 34, the median age being 22 years.

In order that a satisfactory degree of interest might be sustained throughout the duration of the study, all students were told that they would be given their scores upon completion of testing in profile form. In fact, most subjects received scores on those tests completed during the first two class periods at the beginning of the third period. It was thought that additional motivation might be provided if the temporal interval between taking the tests and receiving the scores was not too long.

The Factor Analysis

The matrix of test intercorrelations (all product-moment) presented in Table 2 was factor-analyzed by Thurstone's centroid method in the usual manner with one minor exception (13). In the reflection of signs the criterion was that used by the workers in several of the psychological research units of the United States Army Air Forces during World War II. The algebraic sum of a column, with the diagonal entry disregarded, was employed instead of the mere number of negative signs appearing in a column. This procedure not only tends to guarantee positive sums but also appears to approximate more closely the maximizing of table totals than does the criterion involving number of negative signs.

Because of marked discrepancies between obtained communalities in the first set of centroid extractions and the estimated communalities in the diagonals, a second set of extractions was required. Following the second extraction (of seven centroid factors) the obtained communality of no test differed more than $|.07|$ from the second estimated communality.

The criterion employed for cessation of extraction of the centroid factors was also that used by workers in the psychological research units of the AAF; namely, that factoring should not

TABLE 2

*Product-moment Correlation Coefficients among Fourteen Test Variables**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1. Verbal Comprehension	—	254	035	−045	075	453	182	−047	060	−015	145	−002	028	067
2. General Reasoning	254	—	183	132	285	241	370	135	233	220	372	308	152	271
3. Numerical Operations	035	183	—	201	−029	−038	261	181	088	120	−072	−010	−075	−029
4. Perceptual Speed	−045	132	201	—	254	097	152	517	285	216	305	277	148	354
5. Spatial Orientation	075	285	−029	254	—	317	247	267	405	367	609	467	362	433
6. [Verbal] Completion	453	241	−038	097	317	—	227	188	266	193	402	244	227	342
7. Number Series	182	370	261	152	247	227	—	216	284	239	289	310	303	251
8. Identical Forms	−047	135	181	517	267	188	216	—	293	232	332	265	237	372
9. Cubes	060	233	088	285	405	266	284	293	—	355	396	349	227	367
10. Flags	−015	220	120	216	367	193	239	232	355	—	347	375	313	391
11. Spatial Visualization	145	372	−072	305	609	402	289	332	396	347	—	500	370	500
12. Punched Holes	−002	308	−010	277	467	244	310	265	349	375	500	—	328	541
13. Pattern Analogies	028	152	−075	148	362	227	303	237	227	313	370	328	—	302
14. Form Board	067	271	−029	354	433	342	251	372	367	391	500	541	302	—

* Decimal points omitted.

cease until the product of the two highest factor loadings is at least less than the standard error of the corresponding correlation. Such a criterion tends to yield a greater number of factors than do most other criteria. The rationale underlying this less stringent criterion is that the maximum contribution which the factor makes to the scalar product of two test vectors, or to the correlation between two tests, is no greater than the chance relationship expressed by the standard error of the correlation coefficient.

Following the completion of a set of trial rotations, it was considered advisable to extract two more centroid factors as an aid to further rotations. It was known that probably only six factors would be meaningfully identified. However, previous experience has indicated that use of additional centroid axes in the rotation process frequently brings about, more readily, a psychologically meaningful solution. The superfluous factors eventually appear as mere residuals (factors containing insignificant amounts of communality) to which no interpretation can be dependably given. Moreover, the presence of residual factors seldom interferes at the conclusion of the rotation procedure with the interpretation of those principal factors which account for most of the common-factor variance.

Fifty-six rotations of pairs of axes were required to satisfy Thurstone's criteria of positive manifold and simple structure. Each rotation was achieved graphically according to the method devised by Zimmerman (19). In general the structure determined the direction and magnitude of each new rotation. Information concerning the content of tests was put to use only toward the end of the rotation procedure when minor adjustments were made. In view of the large number of rotations the differences between the communalities of centroid factors and final rotated factors were negligible, the largest two discrepancies being .017 and .013. An orthogonal reference frame appeared to suffice for the interpretation of the factors. The final rotated factor loadings are shown in Table 4.

Interpretation of Factors

Inspection of the final rotated factor loadings in Table 4 reveals that on the whole the criteria of positive manifold and

TABLE 3

*Centroid Factor Loadings and Communalities**

Test	I	II	III	IV	V	VI	VII	VIII	IX	h^2
1. Verbal Comprehension	234	−540	293	−238	−052	−033	103	076	−117	523
2. General Reasoning	502	−300	057	196	168	−133	−085	−038	−063	442
3. Numerical Operations	173	−258	−515	100	128	−127	134	153	−198	485
4. Perceptual Speed	477	228	−363	−259	075	−205	−091	069	081	546
5. Spatial Orientation	648	213	209	150	−133	−127	121	−076	054	589
6. [Verbal] Completion	519	−267	316	−304	−174	111	134	−031	−063	598
7. Number Series	531	−276	−121	259	−093	089	−182	122	−117	518
8. Identical Forms	519	204	−340	−275	−058	−066	−134	069	063	537
9. Cubes	563	093	−071	029	−052	061	152	−099	027	372
10. Flags	527	162	−040	175	081	122	132	099	081	407
11. Spatial Visualization	726	180	306	033	−070	−233	−043	−115	−018	729
12. Punched Holes	630	228	141	129	203	161	−115	−084	−072	578
13. Pattern Analogies	461	159	094	108	−232	172	−128	166	072	392
14. Form Board	662	225	108	−142	203	166	−043	−107	−090	611

* Decima points omitted.

TABLE 4

*Final Rotated Factor Loadings and Communalities**

Test	I (V_z)	II (V)	III (N)	IV (R)	V (V_1)	VI (V_2)	VII (S)	VIII (P)	IX (Res.)	h^2†
1. Verbal Comprehension	−01	70	09	17	−01	03	00	−12	−06	540
2. General Reasoning	24	19	21	54	00	02	10	02	01	437
3. Numerical Operations	17	−03	64	17	−04	01	−04	14	00	495
4. Perceptual Speed	24	−03	22	01	−01	00	07	66	−01	553
5. Spatial Orientation	42	07	−04	20	00	−02	58	16	−04	582
6. [Verbal] Completion	20	66	−02	07	07	19	28	06	05	610
7. Number Series	16	12	34	42	39	06	17	03	05	515
8. Identical Forms	24	03	20	−03	15	06	11	64	06	550
9. Cubes	20	06	13	14	01	18	43	24	10	363
10. Flags	15	−08	11	21	04	26	44	20	13	401
11. Spatial Visualization	62	20	06	25	02	06	44	20	03	722
12. Punched Holes	52	−07	−01	28	08	36	25	14	−01	569
13. Pattern Analogies	24	02	−06	09	41	08	34	15	−08	391
14. Form Board	52	10	01	13	00	43	22	28	02	610

* Decimal points omitted.
† Communalities based on rotated factor loadings expressed to three decimal places.

simple structure have been fulfilled. Six rotated factors were meaningfully identified as visualization (Vz), verbality (V), numerical facility (N), general reasoning (R), spatial-relations (S), and perceptual speed (P). Two other factors (V_1 and V_2) appeared that could not be satisfactorily defined, although their weights in certain tests were suggestive of possible interpretations. A ninth factor turned out as a residual with loadings ranging from −.08 to +.13.

Inasmuch as the primary purpose of the study centered about the investigation of the factors of spatial relations and visualization, the discussion relating to the identification and meaning of the other four factors will be kept to a minimum. The factors, V, N, and P are actually doublets. However, since the factorial content of the pairs of tests weighted in these three factors was well known in advance of their inclusion within the battery, there is little reason to doubt the correctness of the identification given.

It should be pointed out that the major loadings in some tests describing these three factors tended to be somewhat smaller than those reported in other studies or in manuals. This is due to the fact that many of the tests were shortened in order that they might be given within the time period available for testing.[1] However, in view of the size of the sample (N = 360), loadings of .35 or greater are probably indicative of the presence of a significant amount of variance in a factor.

Somewhat greater attention should probably be given to the interpretation of the factor R. Two tests, *General Reasoning* and

[1] It is possible, however, to estimate what the loadings of these three factors, as well as the loadings of the other factors, would be if the tests were not shortened (9). When a test is homogeneously changed in length the new factor loadings may be estimated by the formula

$$k_{mn} = k_{m1}\sqrt{\frac{n}{1 + (n-1)r_{11}}},$$

where n = number of times the test has been lengthened, or the ratio of the length of the new form to the original form;
k_{m1} = loading of factor m in the original, or unlengthened test 1;
k_{mn} = loading of factor m in the lengthened, or new, form of the test;
r_{11} = reliability of the unlengthened test.
If the shortened experimental forms of tests (1), (2), (3), (4), (5), and (11) are considered to be extended to their original length, the corrected loadings in the principal factor in each test are estimated to be respectively, .712, .564, .657, .673, .587, and .631, compared to the obtained loadings of .698, .537, .642, .664, .578, and .619 (which are rounded to two figures in Table 4). The assumption is made in the speed tests that the number of items completed per unit time remains constant.

Number Series, are loaded in this factor to the extent of .54 and .42, respectively. In view of the small number of items contained in the shortened form of the first test (fourteen in all) and of the consequent limitation imposed on the reliability of the test, the magnitude of first loading is substantial. Although the factor may be tentatively described as relating to some type of reasoning function, it is not clearly defined. That it may represent an ability to grasp the essential steps involved in the solution of problems presented in quantitative or symbolic terms appears to be a plausible interpretation.

Interesting to note is the fact that factor V_1 is loaded .39 and .42 in the two tests *Number Series* and *Pattern Analogies*, respectively. A highly speculative interpretation would suggest that this factor may be that of induction previously identified by Thurstone (16). When the possible existence of an induction factor is taken into account along with the fact that the test of *Pattern Analogies* received an insignificant loading of .09 in the factor R, it appears even more plausible that the factor R may represent an ability to diagnose a problem expressed in quantitative terms. If the interpretation of the R factor is correct, a significant finding is that a test (*General Reasoning*) can be constructed to measure quantitative thinking without the introduction of substantial amounts of variance in the numerical factor.

Examination of the loadings for the final rotated factors I and VII in Table 4 reveals positive, though not conclusive, evidence for the existence of two reference variables which may be meaningfully identified as spatial-relations and visualization. In short, the two hypotheses as set forth are, in the main, upheld—at least to the extent that the factorial composition of the two groups of selected tests differs.

In the following list of four tests, the first three of which were selected to test the hypothesis relating to the psychological processes involved in visualization, loadings of .35 or higher in all rotated factors including I (Vz), VII (S), and VI (V_2) may be summarized as follows:

Tests	Factor I (Vz)	Other Factors
(11) Spatial Visualization	.62	.44S
(12) Punched Holes	.52	.36V_2 (.25S)
(14) Form Board	.52	.43V_2 (.22S)
(5) Spatial Orientation	.42	.58S

In view of the presence of weights of .52 or higher in three tests *Spatial Visualization*, *Punched Holes*, and *Form Board*, (all three of which made up the group intended to represent a measure of visualization), factor I can be identified as *visualization*, even though the test of *Spatial Visualization* is loaded to the extent of .44 in factor VII (S). That the spatial-relations and visualization abilities may be required in one or more tests in either of the two groups of tests inserted in the battery was mentioned previously as a definite possibility.

After taking the test of *Spatial Visualization*, many of the subjects reported that in addition to manipulating mentally the stimulus figure (an alarm clock) into the final position called for by the verbal directions, they also related the location of various parts of the stimulus object (hands, numerals, top, base, winding and setting mechanisms of the clock) to the location of corresponding parts of one or more response figures (five alarm clocks in different positions). In the easier items which required only one manipulation the role of spatial cues is undoubtedly important. On the other hand, in those items requiring two or three movements of the clock, it would appear that a greater dependence was placed upon manipulations of the clock; in fact, in the most difficult items variance associated with reasoning, verbal, and memory factors would possibly be important. However, only four items requiring a sequence of three movements were scored. Nevertheless, a small, though perhaps insignificant, loading of .25 appeared in the R factor. In short, the influence of the range of difficulty of items upon the factorial content of a test may be substantial, as a previous study has shown (6).

In two other tests, *Punched Holes* and *Form Board*, which were weighted heavily in the visualization factor, small loadings of .25 and .22, respectively, appear in the factor to be identified as spatial-relations. More important, however, are the corresponding loadings of .36 and .43 in a factor V_1. Although not amenable to a dependable identification, this factor may be associated with the drawing (filling in) response required of the examinees. Despite their relatively high saturations in the visualization factor, these two tests appear to involve additional unknown factors.

The visualization factor loading of .42 in the test of *Spatial*

Orientation, which was chosen to represent a measure of the spatial-relations factor, is probably indicative of the use by some of the examinees of visualization. Introspective reports from subjects differed as to the technique used in working the items. The variance representing visualization ability may be attributed to the tendency of several examinees mentally to manipulate the boat, as if it were a small toy, up and down and/or left or right and to imagine concomitant changes in the scenery. Many of the subjects reported that they did not place themselves within the boat, but viewed the boat and scenery as if they were on a stationary platform some distance to the rear of the boat. One subject said that he pretended to be playing with a toy boat in a pond and to be sighting along the prow of the boat as a means of observing shifts in background scenery while he moved the boat with his hand to the right or left and/or up or down.

On the other hand, many, if not most, of the subjects pretending actually to be inside the boat, and using the prow as the guide, noted changes in background views with reference to corresponding motions of the boat. Although the test of *Spatial Orientation* appears to be weighted in both spatial-relations and visualization factors, it does seem to represent best a measure of spatial relations or spatial orientation and to vindicate its inclusion with other tests in the battery which were selected to bring out the spatial factor.

In the following list of five tests, the first three of which were chosen to yield evidence regarding the second hypothesis, loadings of .34 or higher were found in rotated factors VII (S), I (V_Z), and V (V_1):

Tests	Factor VII (S)	Other Factors
(5) Spatial Orientation	.58	.42 (V_Z)
(10) Flags	.44	(.15 V_Z)
(11) Cubes	.43	(.20V_Z)
(12) Spatial Visualization	.44	.62V_Z
(13) Pattern Analogies	.34	.41 V_1(.24 V_Z)

The magnitude of the weights in factor VII for the tests of *Spatial Orientation*, *Flags*, and *Cubes* indicates that identification of the factor as spatial relations is psychologically meaningful. Despite the substantial loading of the visualization factor

in the test of *Spatial Orientation*—a fact which has been rationalized previously—the first hypothesis regarding the psychological nature of spatial relations appears to have been upheld.

Of passing interest is the loading of .34 in the spatial relations factor appearing in the test of *Pattern Analogies*. In this factorially complex test, the presence of variance in the spatial-relations factor may have been due to the role of those changes in the design of complex figures, or patterns, which depended upon a rule involving the spatial order of parts. In the more difficult items of complex design it was usually helpful, if not necessary, to give specific attention to the spatial organization of the various geometric properties within each of the patterns appearing in the row.

A second source for possible variance in the spatial-relations factor was that of the format of each item. Pattern A and pattern B, which stood in a left-right order on the page, corresponded to the order of pattern C and one of the five alternative responses. Having been exposed to spatial tests administered earlier, many of the subjects may have transferred techniques previously learned in solving other items to the task required in the test of *Punched Holes*. Thus, the influence of mental set may have been one important reason for the appearance of the loading in the spatial-relations factor.

The results of the factor analysis seem to indicate that, in the main, the two hypotheses have been upheld. Two of the final rotated factors may be readily interpreted in terms of their weights in two groups of tests as representing the spatial-relations and visualization abilities that were hypothesized. However, the number of tests does not appear to be large enough to determine with confidence whether the abilities may be correlated to some degree.

Much needed, indeed, are other studies to yield further evidence regarding the tenability of these two hypotheses. Although two recent empirical investigations (1, 14) have indicated that similar primary factors are obtained when the same, or nearly the same, batteries of tests are administered to groups chosen under different selective conditions, it is urged that other homogeneous samples in which such variables as age, level of educational attainment, occupational classifica-

tion, and sex membership are systematically varied be employed to test the validity of the two hypotheses. Other hypotheses should be formulated regarding the psychological nature of the spatial domain and subjected to verification through use of specially devised tests and of other tests of known factorial composition. It is hoped that following more extensive research in the area of space and visualization relatively pure tests can be constructed[4] to measure the abilities identified and that such tests can be used with others of demonstrated merit to improve materially the degree of accuracy with which numerous complex criteria can be predicted.

Summary

The primary purpose of the study was to test the tenability of two hypotheses regarding the psychological nature of spatial-relations and visualization factors. A secondary purpose was to seek to identify certain factors found in the AAF investigations with certain of Thurstone's primary abilities. Within a battery of fourteen tests, two groups of tests (three tests in each group) were included which appeared to reflect differences in the psychological processes associated with the spatial-relations and visualization abilities. In addition to the six tests expressly incorporated within the battery to yield evidence regarding the validity of the hypotheses, eight reference tests of fairly well-known factorial content were included to aid in the identification of variance found in the six tests and to answer questions of identity of the Thurstone and AAF factors.

Positive evidence for the hypotheses was to be considered attained if the two groups of tests defined separate factors and if none of the other eight tests was substantially weighted in factors unique to either group of tests. Moreover, none of the three tests in one group should contain large amounts of variance in common with tests of the other group except to the extent that a given test might consist of items that reflected the presence of that factor which was defined in the main by

[4] Even if pure tests cannot be constructed for all factors identified in the spatial realm, means are available for attaining estimates of univocal factor scores through use of suppression tests (8).

tests of the other group. If a test did appear in one group that contained variance in the factor associated primarily with tests of the other group, a satisfactory rationalization of this finding would be required.

Product-moment correlations computed from sets of scores of 360 students in the introductory course in psychology at Rutgers University were factored by Thurstone's centroid method. Eight of these factors were rotated by graphical means to positions satisfying the criteria of positive manifold and simple structure.

In the orthogonal system six factors were identified as verbal comprehension, numerical facility, perceptual speed, reasoning, visualization, and spatial relations. In the main, the variances associated with factors identified as spatial relations and visualization were confined to the respective groups of tests initially placed within the battery to bring out the factors. In only one test in each group of three tests were substantial amounts of variance found in both the visualization and spatial-relations factors, although the larger portion of variance was in the factor common to the group in which that test appeared.

The presence of variance in these two factors was rationalized for each of the tests. Introspective reports of the subjects revealed that in many items the psychological processes used involved both spatial-relations and visualization abilities as described in the hypotheses. The range of difficulty level of test items in one test also appeared to be an important reason for the appearance of two factors.

In short, it may be concluded that the two hypotheses regarding the psychological nature of visualization and spatial relations were confirmed. However, other research projects need to be carried out with a variety of samples before a dependable generalization can be made regarding the nature of these two abilities. Since there is some evidence of still other spatial abilities (3), some or all of which may be correlated, it is recommended that a conscientious attempt be made to formulate in operational terms new hypotheses and that new tests, having been constructed in harmony with the hypotheses, be factor analyzed along with other tests of established factorial content. Once the area of space has been dependably and adequately mapped, attention can be directed toward building tests approximating pure measures of the identified abilities.

REFERENCES

1. Dudek, F. J. "The Dependence of Factorial Composition of Aptitude Tests Upon Population Differences Among Pilot Trainees. I. The Isolation of Factors." EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, VIII (1948), 613–633.
2. Fruchter, B. "The Nature of Verbal Fluency." EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, VIII (1948), 33–47.
3. Guilford, J. P. (Ed.) *Printed Classification Tests*, Report No. 5. *Army Air Forces Aviation Psychology Program Research Reports*. Washington, D. C.: U. S. Government Printing Office, 1947.
4. Guilford, J. P. "Factor Analysis in a Test-Development Program." *Psychological Review*, V (1948), 79–94.
5. Guilford, J. P. "Some Lessons from Aviation Psychology." *The American Psychologist*, III (1948), 3–11.
6. Guilford, J. P. "The Difficulty of a Test and its Factor Composition." *Psychometrika*, VI (1941), 67–77.
7. Guilford, J. P. "The Discovery of Aptitude and Achievement Variables." *Science*, CVI (1947), 279–282.
8. Guilford, J. P. and Michael, W. B. "Approaches to Univocal Factor Scores." *Psychometrika*, XIII (1948), 1–22.
9. Guilford, J. P. and Michael, W. B. "Estimates of Factor Loadings When a Test is Homogeneously Changed in Length." *Psychometrika*, (to be printed).
10. Guilford, J. P. and Zimmerman, W. S. "Some AAF Findings Concerning Aptitude Factors." *Occupations*, XXVI (1947), 154–159.
11. Guilford, J. P. and Zimmerman, W. S. *The Guilford-Zimmerman Aptitude Survey*. Beverly Hills, Calif.: Sheridan Supply Company, 1947.
12. Guilford, J. P. and Zimmerman, W. S. "The Guilford-Zimmerman Aptitude Survey." *Journal of Applied Psychology*, XXXII (1948), 24–34.
13. Kelley, T. L. *Crossroads in the Mind of Man*. Stanford University: Stanford Univ. Press, 1928.
14. Michael, W. B. "Factor Analyses of Tests and Criteria: A Comparative Study of Two AAF Cadet Pilot Populations." *Psychological Monographs: General and Applied*, 1949, No. 298.
15. Thurstone, L. L. *Multiple Factor Analysis*. Chicago: University of Chicago Press, 1947.
16. Thurstone, L. L. "Primary Mental Abilities." *Psychometric Monographs*, No. 1. Chicago: University of Chicago Press, 1938.
17. Thurstone, L. L. "The Perceptual Factor." *Psychometrika*, III (1938), 1–17.

18. Thurstone, L. L. and Thurstone, T. G. "Factorial Studies of Intelligence." *Psychometric Monographs*, No. 2. Chicago: University of Chicago Press, 1941.
19. Zimmerman, W. S. "A Simple Graphical Method for Orthogonal Rotation of Axes." *Psychometrika*, XI (1946), 51–55.
20. Zimmerman, W. S. "Isolation, Definition, and Measurement of Spatial-Visualization Abilities." Ph.D. dissertation, University of Southern California, 1949.

ON THE USE OF INTERACTIONS AS "ERROR TERMS" IN THE ANALYSIS OF VARIANCE[1]

ALLEN L. EDWARDS
University of Washington

I.

MANY psychological and educational experiments are concerned with two or more variables, each of which may be varied in two or more ways. When the variables are studied in all possible combinations in the same experiment, the experiment is said to be of *factorial* design.[2] As an example, let us take an experiment in which three variables are involved, A, B, and C. Suppose that A is varied in three ways, B is varied in two ways, and C is varied in four ways. Then we shall have (3) (2) (4) = 24 combinations of variables, each combination corresponding to a particular experimental condition. One replication of the experiment will thus require 24 observations and the 23 degrees of freedom available with one replication would be allocated in the following way:

Sum of squares		*df*
Main variables	A	2
	B	1
	C	3
First order interactions:	$A \times B$	2
	$A \times C$	6
	$B \times C$	3
Second order interactions:	$A \times B \times C$	6

If 240 subjects were available, then 10 could be assigned at random to each of the 24 experimental conditions. We would thus have 9 degrees of freedom within each of the experimental conditions or (9) (24) = 216 degrees of freedom for the varia-

[1] This paper is based upon a section of a manuscript which deals more extensively with problems of experimental design in psychological and educational research. I should like to acknowledge that I have incorporated into this paper the suggestions of Dr. Paul Horst, who served as a technical consultant on the manuscript.

[2] It is assumed that the reader is familiar with the treatment of the analysis of variance as given, for example, by Lindquist (6), McNemar (7), or Snedecor (8).