6483

16. Pace, Robert C. "Opinion and Action: A Study in Validity of Attitude Measurement." Educational and Psychological Measurement, X (1950), 411–419.
17. Preston, Malcolm G. "Note on the Reliability and the Validity of the Group Judgment." *Journal of Experimental Psychology*, XXII (1938), 462–471.
18. Ryans, David G. "A Note on Methods of Test Validation." *Journal of Educational Psychology*, XXX (1939), 315–319.
19. Spencer, Douglas. "Frankness of Subjects on Personality Measures." *Journal of Educational Psychology*, XXIX (1938), 26–35.
20. Thouless, R. H. "Test Unreliability and Functional Fluctuation." *British Journal of Psychology*, XXVI (1935–1936), 325–343.
21. Weiss, L. A. "Rating Scales, with Special Reference to the Field of Child Development," *Psychological Bulletin*, XXX (1933), 185–208.
22. Weinland, J. D. "Use of Rating Scales and Personal Inventories to Check Each Other." *Journal of Applied Psychology*, XXXII (1948), 631–635.
23. Windle, Charles. "Test-Retest Effect on Personality Questionnaires." Educational and Psychological Measurement, XIV (1954), 617–633.
24. Young, H. H., Holtzman, W. H. and Bryant, N. D. "Effects of Item Context and Order on Personality Ratings." Educational and Psychological Measurement, XIV (1954), 499–517.

6483
DAVI 41

A FACTOR ANALYSIS OF THE WECHSLER-BELLEVUE SCALE[1]

PAUL C. DAVIS
University of Southern California[2]

Since its introduction in 1939, the Wechsler-Bellevue Scale (30) had gained widespread popularity as a clinical instrument and as a test for the individual measurement of intelligence. During this period, only a few of the hundreds of articles relating to the W-B (21, 22) have dealt with the question of what the scale measures. It is to this problem that the present research was directed.

Although the title of the scale itself implies that it measures "intelligence", it is quite generally recognized that intelligence is not the unitary trait that the term might suggest. Wechsler himself indicates the probable complexity of intelligence as he has attempted to measure it, by calling it an "aggregate because it is composed of elements or abilities" (30 p. 3). Recognizing this complexity, Balinsky (1), Goldfarb (15), Birren (2, 3), Cohen (6), and Simkin (23) have all reported analyses designed to isolate factors of human ability in which the subtests of the W-B have significant variance. Careful study of these researches convinced the present author that a further step was necessary, namely, that of relating Wechsler factor findings to those factors well established by other research. In order to accomplish this step, it appeared necessary to include in the intercorrelation matrix not only the W-B subtests but also reference tests of known factor content. In this frame of reference, the two-fold purpose of the present study might be stated as (a) determination of the factor content of the W-B in terms of known factors, and (b) the inter-

[1] This paper is based on a dissertation submitted to the Department of Psychology of the University of Southern California (9) in partial fulfillment of the requirements for the Ph.D. degree. The writer wishes to express special appreciation to Drs. L. G. Humphreys and Paul Horst who were most helpful in the planning and early stages of the study, and to Dr. J. P. Guilford, chairman of the doctoral committee, under whose generous and patient supervision the research was completed.

[2] Now at Roberts Wesleyan College, North Chili, New York.

pretation of factor loadings and descriptions in such a way that they would be meaningful to those using the scale, many of whom have little or no knowledge of factor analysis.

Method

Design. Examination of even a few studies that have attempted to analyze intelligence or intelligence test scores (3, 4, 5, 19, 26, 32) reveals some twenty or more factors with different names and descriptions. This situation called for careful rational analysis of the W-B for the purpose of guidance in the selection of reference tests that would assure identification of the significant variances in the W-B Scale. On the basis of logical considerations and review of previous factor analysis findings significant W-B variances in the following factors were hypothesized: (a) verbal comprehension, (b) numerical facility, (c) visualization, (d) perceptual speed, (e) reasoning, (f) mechanical knowledge. The possibility was also considered that the W-B might contain variance in a memory factor and a space relations factor, but no reference tests were included for these factors. Results here reported suggest that the omission of a space relations reference variable was unfortunate, resulting probably in the distribution of space relations variance among several other factors or its appearance as specific variance.

Other hypotheses were formulated, having to do with method of administration and type of test. These hypotheses accounted for the inclusion of certain variables in the battery, but limitations of space do not allow their discussion in the present report.

In line with the above hypotheses, the following reference tests were planned and prepared.

a. two verbal comprehension tests (variables 9 and 10)
b. one numerical facility test (variable 4)
c. one visualization test (variable 6)
d. one perceptual speed test (variable 3)
e. one reasoning test (variable 5)
f. one mechanical knowledge test (variable 7)

Description of Variables. The final design included the 24 variables that are listed below. Accompanying numbers identify the variables in tables and discussions that follow.

1. Age
2. Otis Beta
3. Perceptual Speed
4. Numerical Operations
5. Arithmetic Reasoning
6. Mechanical Principles
7. Mechanical Information
8. Science Information
9. Vocabulary-Recognition
10. Vocabulary-Recall
11. W-B Group Information
12. W-B Group Comprehension
13. W-B Group Similarities
14. W-B Information
15. W-B Comprehension
16. W-B Digit Span
17. W-B Arithmetic
18. W-B Similarities
19. W-B Vocabulary
20. W-B Picture Arrangement
21. W-B Picture Completion
22. W-B Block Design
23. W-B Object Assembly
24. W-B Digit Symbol

Age was computed to the nearest month. The scale was reversed so that high score is associated with low age. *Otis Beta* scores were obtained from school records. The *perceptual speed* test was designed by the writer[3] to measure the well-known perceptual speed factor. The test items consist of key figures constructed by mutilating capital letters of type known to printers as "Wedding Text", each key figure followed by five alternative figures similarly constructed. One or more of the five is exactly like the key figure. On an overprinted IBM answer sheet, the subject marked the figures identical to the key. The *numerical operations* test consisted of simple addition, subtraction, multiplication, and division, items similar to those in the AAF test of the same name (18). The subject selected the one correct answer from five alternatives. The *arithmetic reasoning* test was constructed by the writer along

[3] The idea for this test was suggested by Dr. L. G. Humphreys.

traditional lines. Written problems required considerable reasoning but a simple level of numerical skill. Five alternative answers were provided. The *mechanical principles* test was an adaptation of the AAF test of the same name (18) consisting of pictorial presentation of everyday mechanical problems. The *science information* was a verbal test of present day scientific facts, not designed as a factor-reference test. The two vocabulary tests were constructed by the writer to be equivalent in content and difficulty. Items were listed in ascending order of difficulty level as determined by Thorndike's Word List (25). In the *vocabulary recognition* test, the subject selected the one of five alternative words with meaning most nearly equivalent to that of the key word. The *vocabulary recall* test consisted of a list of words for which the subject was required to write definitions. Variables 11, 12, and 13 were simply printed, group adaptations of W-B, Form II tests of the same names. Variables 14 through 24 were the subtests of the W-B Scale, Form I, administered and scored in prescribed fashion (30).

Subjects and Administration. The sample tested was composed of 356 students in Hamilton Junior High School in Seattle, Washington.[4] Students were all in grade 8A and ranged in age from 12-5 to 16-10. The sample came from upper-lower and lower-middle class homes and was approximately evenly divided between the sexes. Scores on variables 1 through 13 were available for all students, while the W-B (variables 14–24) was administered to only 202 of the total. Group tests were administered to pupils by sections of 75 to 100 under carefully standardized conditions. Variables 3–5 were highly speeded, while variables 6–10 were administered so that most students had the opportunity of trying all items.

Factor Analysis

After raw scores were transformed into stanine (18) scores and intercorrelations computed, Thurstone's complete centroid method was employed for the extraction of factors.[5]

[4] Thanks are hereby extended to the Seattle Schools and to Dr. Lyle Stewart, Assistant Superintendent, for making this testing program possible.

[5] Three tables presenting product-moment correlations of the variables, centroid factor loadings, and rotated factor loadings have been deposited with the American Documentation Institute. Order Document No. 4703 from American Documentation Institute, 1719 N. St. N.W., Washington 6, D. C., remitting $1.25 for 35 mm microfilm or $1.25 for photoprints readable without optical aid. Copies of tables are also available from the author without cost.

Criteria suggested by Thurstone (27) and others were essentialy met after extraction of the tenth factor, so extraction was continued to eleven factors only.

Orthogonal rotation of reference axes was carried out by means of the graphic method developed by Zimmerman (33). In view of the design of the study, the reference tests (variables 3, 4, 5, 6, 7, 9, 10) were utilized as sign-posts in the rotation. Thurstone's criteria of simple structure and positive manifold along with the principle of psychological meaning, were employed as guides. A total of 101 rotations resulted in what appeared to be a stable solution, with ten significant factors and one that is apparently a genuine residual.

Results

Interpretation of Factors. Since the standard error of a factor loading is not known, decision as to significance of loadings is to some extent arbitrary. Loadings below approximately .30 have not been interpreted in the present study although some smaller loadings are suggestive.

Factor A, verbal comprehension.

10. Vocabulary-recall	.795
9. Vocabulary-recognition	.751
11. W-B Group Information	.609
19. W-B Vocabulary	.601
8. Science information	.572
2. Otis Beta	.538
14. W-B Information	.423
12. W-B Group Comprehension	.375
18. W-B Similarities	.310

This is quite clearly the verbal comprehension factor that has been reported in many analyses (12). Aside from vocabulary tests, information tests are most heavily loaded on this factor. This is in line with the AAF findings which reported five analyses in which information tests had appreciable loadings on a verbal factor. Balinsky (1) reported the W-B Information test to have a verbal factor loading of .49. Appearance of the Otis Beta on this factor has considerable precedent according to French (12). Science information and the W-B Group Information and Similarities loadings are easily rationalized on the basis of the verbal comprehension and facility required.

Factor B, visualization.

6. Mechanical principles	.641
22. W-B Block Design	.439
21. W-B Picture Completion	.384
23. W-B Object Assembly	.338

This factor is identified chiefly by the moderately high loading of the mechanical principles test. This test was the reference variable for the factor, having been identified with it repeatedly in the AAF research (18). Since the visualization loading is somewhat larger than has been typically reported for the test, however, it appears possible that the factor is adulterated with certain other variance. The best hypothesis is probably that some space variance has also been included in this factor. The presence of the three W-B tests on the factor is significant and well in line with its interpretation as the visualization factor. Birren (3), Cohen (6), and Balinsky (1) all reported more general factors on which these and other W-B tests appeared, variously described as closure, nonverbal, and performance. Simkin (23) reported a visual-motor factor but his findings were not available for perusal. Guilford's (18) comment that visualization is strongest in tests that present stimuli either pictorially or verbally and require some implicit manipulation tends to confirm the interpretation of factor B as visualization. Particularly in view of the appearance of the Picture Completion test on the factor, the hypothesis is offered that visualization covers not only implicit manipulation (of both actual and symbolic objects) but also the formation of mental images of *eventual relationships* of the objects.

Factor C, numerical facility.

4. Numerical operations	.740
24. W-B Digit Symbol	.524
17. W-B Arithmetic	.359
1. Age	.320
2. Otis Beta	.304
3. Perceptual speed	.303

This factor is identified particularly by the numerical operations test which is similar to those tests that repeatedly identified the numerical factor in the AAF analyses (18). French (12) cites some 35 analyses in which such a factor has emerged. Both the numerical operations and W-B Arithmetic test loadings agree well with previous factor results and interpretations.

The presence of the Digit Symbol test with moderate loading on the numerical factor was a surprise, as was the loading for the perceptual speed test. The appearance of age and Otis Beta scores also calls for some explanation. Attention might first be directed to the typical description of the factor. French (12) says that it involves manipulation of numbers, loadings being roughly proportional to amount of manipulation required. In terms of psychological processes, manipulation certainly has meaning. On the other hand, *numbers* as such are merely stylized symbols that do not necessarily define the boundaries of the psychological task. Coombs' (8) interpretation seems to suggest a similar line of reasoning. In view of this fact as well as other evidence, the hypothesis is offered that the factor called numerical facility or number (12) is best described as to ability to manipulate simple, spatial symbols that are relatively independent (as contrasted to verbal symbols).

The Otis test includes sprinkling of numerical problems. It seems probable, however, that the moderate loadings of this test and of age are mainly functions of selective factors which allow younger and more capable children to progress more rapidly in school.

Factor D, mechanical knowledge.

7. Mechanical information	.436
22. W-B Block Design	.411
6. Mechanical principles	.393
17. W-B Arithmetic	.340
16. W-B Digit Span	.336
8. Science information	.317
5. Arithmetic reasoning	.315

The first three loadings suggest that this is the same factor that was identified repeatedly in AAF analyses and in analyses by Comrey (7), Dudek (10, 11), and Michael (20). These studies as well as the present findings suggest that this is what French (12) calls an *experiential* factor. The various analyses have reported a wide variety of tests on the factor, indicating that it is very broadly based. The presence of information tests and Digit Span suggest that memory is operating in determining the variance on this factor. This is not strange, since memory is certainly important in the retention and use of experience. The results based on this relatively young group suggest that the factor had not yet reached the stage of specificity that was demonstrated with older subjects such as Air Force cadets. This lack of specificity may inhere in part also in the fact that both sexes were included in the present study. Dudek's (11) finding suggests the latter interpretation. On the other hand, a similar outcome would be predicted on the basis of Garrett's (14) hypothesis concerning the differentiating effect of maturation. In any event, factor D appears to be described more accurately by the term *general technical sophistication* than by the term mechanical experience.

Factor E, similarities doublet.

13. W-B Group Similarities	.651
18. W-B Similarities	.470

Behavior of this factor throughout rotation made the conclusion quite compelling that it is true doublet. Very little variance of other tests remained on this factor when rotations were complete. If the two forms of Similarities are truly parallel the appearance of a doublet would be the predicted outcome (27).

Factor F, general reasoning.

17. W-B Arithmetic	.571
5. Arithmetic reasoning	.436
15. W-B Comprehension	.331
13. W-B Group Similarities	.320

On the basis of the first two loadings, this factor was interpreted as the general reasoning factor that was reported in several of the AAF analyses (18). In his analysis of reasoning, Green (16) reported such a factor. His leading test was called Problem Solving and is very similar to the arithmetic reasoning test of the present analysis. Appearance of the Comprehension test is well in line with the interpretation of the factor as reasoning. The absence of the group form (Form II) of Comprehension may be rationalized in part in terms of the necessity of multiple response required (see discussion of factor G). Appearance of the group form of Similarities with larger loading than that of the individual form, (.188) may be an artifact related to the abnormally high communality of the former.

Factor G, fluency.

12. W-B Group Comprehension	.625
16. W-B Digit Span	.369
24. W-B Digit Symbol	.304

It will be noted that not only the above variables but also all those with loadings above .20 (Comprehension, Picture Completion, Vocabulary, and group Similarities) are W-B tests. All but Digit Symbol require *verbal response*. Careful analysis of the list reveals one especially significant fact, that the loading of the group Comprehension (Form II) test was high on the factor while the loading of Comprehension, Form I, was low. Examination of the two forms revealed one major difference. In Form I, only one item requires multiple answers for full credit. In Form II, however, five of the ten items demand multiple answers to a single question. Since Form II has the outstanding loading on the factor, this difference appears to be significant. Interpretation of this fact in the light of findings of Fruchter (13) and Taylor (24) and the summary of French (12) led to the designation of the factor as fluency. Once again, as in the case of Factor D, the results do not justify restricting the term by the use of a modifying adjective. The factor seems to identify a tendency to be responsive and uninhibited, with consequent free flow of responses to stimulation. This may well be mainly what Wechsler (28, 29, 31) calls non-intellective or conative in character.

Factor H, perceptual speed.

3. Perceptual Speed	.515
23. Object Assembly	.415
22. Block Design	.385
13. W-B Group Similarities	.378
24. W-B Digit Symbol	.366
4. Numerical Operations	.311
9. Vocabulary-recognition	.310

Many analyses have resulted in the identification of a factor of perceptual speed, involving, as French (12) has indicated, the speedy identification of a given configuration within a mass of distracting material. Tests 3, 23, and 22 quite clearly involve perceptual speed in this sense. As Green (16) has pointed out also, perceptual speed tests must have a low absolute level of difficulty. The three tests referred to also meet this criterion. Beyond these accepted characteristics of the factor, it appears possible that the present analysis has identified another. The Object Assembly test came out with approximately equal loadings on Visualization and Perceptual Speed. It is at least hypothesized that the variance in perceptual speed has to do with rapid identification of shapes that will fit together, while the variance in visualization has to do with the eventual whole more than with the parts. Experience in administering this test reveals at least two distinctly different approaches to the task in spite of the visualization oriented directions. It is possible that the break is along the lines suggested above.

It appears probable that again the loading on variable 13 is artifactually high as a result of the inflated communality of the variable. On the other hand, the loadings of variables 13, 4, and 9 above present some interpretative difficulty. Perhaps, as Fruchter (13) suggests, loadings of such tests reflect variance in reading speed. It would be of interest to know the extent to which reading speed is dependent upon the factor of perceptual speed.

Factor J, eduction of conceptual relations.

20. W-B Picture Arrangement	.494
2. Otis Beta	.474
1. Age	.396
10. Vocabulary-recall	.372

In spite of the apparent lack of psychological meaning in the above grouping of tests, factor J resisted rotational efforts to modify the essential picture. One clue that led to the above identification was the fact that the Thurstones (26) reported an appreciable loading (−.22) on age for a factor they called induction. Coupled with this, was the logical hypothesis that the Picture Arrangement test involves reasoning of some kind. Green's (16) study of reasoning provided the name for the factor although the present study did not make possible the fine distinctions that Green obtained. The appearance of the Otis Beta on the factor is rationalized by the obvious importance of reasoning in intelligence as it is typically measured and by the fact that the Thurstones, in the study cited above, reported a loading of .35 for mental age on the same induction factor. The presence of variable 10 seems to be explained by the fact that the test was difficult and required free type of response. The seeking for definitions for words not fully known can logically be thought of as a sort of semantic eductive process.

It might be of interest to note that the present study tends to confirm the opinion of others (11, 18, p. 122) that factors of human ability in reasoning do not necessarily break along the classical lines of logic (i.e. induction and deduction).

Factor K, information.

14. W-B Information	.560
11. W-B Group Information	.439
13. W-B Group Similarities	.437
17. W-B Arithmetic	.318
22. W-B Block Design	.305

Factor K proved the most difficult to identify and rationalize satisfactorily, and for this reason complete discussion of the problem is impossible here. Lines of reasoning can be indicated briefly, however. First, loadings significantly above .30 appear to be in the area of information (discounting somewhat the loading of variable 13, for reasons previously cited). Second, both Balinsky (1) and Goldfarb (15) reported factors associated with awareness of environment or experience. Third, generally (3, 6, 17, 18) the principal variance of information tests has been found to be associated with the verbal comprehension factor. Fourth, the inclusion of parallel forms of the test of information (11 and 14) might well give rise to the appearance of a doublet factor (27).

Consideration along the lines indicated above led to interpretation of the factor as "information" without further specification. In any case, it appears to be a measure of familiarity with facts of the world about, and possibly of familiarity with objects and symbols common to that world, as well.

Discussion

Hypotheses. In part due to the conservativeness of the hypotheses and in part due to the inclusion of reference variables of known factors, all factors hypothesized appeared in the solution. Although space relations and memory variances were suspected in the W-B scale, no variables were included for these factors and no obtained factors could be identified as either of these. This outcome may reflect absence of variance in these factors, but probably in part reflects absence of strong reference variables for them. There is some reason to believe that a certain amount of space variance was confounded with visualization, while in all probability both space and memory are represented in the specific variances of various tests.

In addition to the factors hypothesized, four other factors emerged, at least two of which are interpreted as signalling the presence of hitherto unidentified factor variance chiefly in variables other than reference tests. This outcome indicates that the complexity of the W-B scale was underestimated by the hypotheses. The degree of underestimation is impossible to determine, but it appears certain that at least as many factors as subtests are represented importantly in the W-B scale. Further research, with more complete reference variable coverage, appears to be called for.

Factorial Content of W-B Subtests. As noted earlier in this report, one important objective of the study was to ascertain for practical purposes what the various subtests of the Wechsler scale measure. In this section, each subtest title is followed by a listing of the factors in which appreciable variance appeared, with the *proportion of variance* ascribable to each in the present study. Small variances in other factors have been grouped together as "other". The difference between the total of proportions for a given subtest and unity represents specific and error variance, unidentified otherwise by this study. Variable numbers are indicated.

1. Information (14)

K. Information	.31
A. Verbal Comprehension	.18
C. Numerical Facility	.06
D. Mechanical Knowledge	.05
Other	.10

As Wechsler (30) has suggested, it appears that information makes a unique contribution to the scale. The contribution apparently is chiefly made by the factor called "information". Referring again to factor A, it is of interest to note that only Vocabulary is a better W-B measure of the verbal comprehension factor than is the Information test.

2. Comprehension (15)

F. General Reasoning	.11
A. Verbal Comprehension	.08
G. Fluency	.08
B. Visualization	.07
C. Numerical Facility	.05
Other	.03

Not only is the Comprehension test extremely complex, it is also a relatively poor measure of those factors in which variance appears. From the standpoint of factor content, every factor measured by the scale is better measured by some subtest other than Comprehension. This fact suggests that the one of the six verbal tests in the scale to be omitted in administration might well be Comprehension. In this connection, it is of considerable interest to note the factor picture for the Comprehension test from Form II of the W-B (administered as a group test in this study).

Comprehension (Form II) (12)

G. Fluency	.39
A. Verbal Comprehension	.14
D. Mechanical Knowledge	.06
Other	.14

In contrast to the Form I version, this test is relatively purer factorially and is by far the best single measure of the factor known in this study as fluency, a factor hypothesized to be of the non-intellective type that Wechsler considers so important.

3. Digit Span (16)

G. Fluency	.14
D. Mechanical Knowledge	.11
C. Numerical Facility	.08
Other	.11

Of the Form I tests, Digit Span is the best measure of the fluency factor. In view of this fact and of its brevity and in spite of its limited reliability, Digit Span should probably not be the test omitted from the verbal scale administration. It is of interest to note that the fluency factor is probably not unrelated to the diagnostic value frequently ascribed to the test by clinicians.

4. Arithmetic (17)

F. General Reasoning	.33
C. Numerical Facility	.13
D. Mechanical Knowledge	.12
K. Information	.10
Other	.13

It is interesting that this test, consisting of only a few problems and administered without pencil or paper, proves to be a better measure of general reasoning than is a much longer group arithmetic test. With one third of the total variance associated with the factor, this test is a fairly good measure of general reasoning.

5. Similarities (18)

E. Similarities doublet	.22
A. Verbal Comprehension	.10
K. Information	.08
B. Visualization	.08
Other	.10

Although the nature of the doublet factor cannot be determined further, it is certainly possible that the test makes a unique contribution. Other factors represented are better measured by other tests.

6. Vocabulary (19)

A. Verbal Comprehension	.36
B. Visualization	.09
F. General Reasoning	.07
G. Fluency	.06
Other	.16

Vocabulary lives up to its reputation for being the best single measure of verbal comprehension. It is the purest factor test and one of the most reliable of the scale. Although other tests contribute to the measurement of the verbal comprehension factor, the above facts suggest that Vocabulary should be omitted in administration of the scale only after careful consideration.

7. Picture Arrangement (20)

J. Education of conceptual relations	.24
B. Visualization	.08
Other	.10

One unexpected outcome was the identification of a second reasoning factor by this test, one apparently quite distinct from that measured by arithmetic. This almost unique contribution to the scale points up the interpretative importance of the P-A test score (with due regard, of course, to matters of reliability).

8. Picture Completion (21)

B. Visualization	.15
G. Fluency	.08
J. Education of Conceptual Relations	.07
K. Information	.07
Other	.11

This is the one performance test that apparently measures only those common factors that are better measured by some other test of the scale. In this respect it is probably the least useful of the performance tests. On the other hand, there is probably considerable specific variance that was not identified by this analysis.

9. Block Design (22)

B. Visualization	.19
D. Mechanical Knowledge	.17
H. Perceptual Speed	.15
K. Information	.09
J. Education of conceptual Relations	.05
Other	.05

Block Design had the highest communality of all the performance tests, leaving, in all probability, very little specific variance. Although quite complex, it contributes significantly to the measurement of its three leading factors.

10. Object Assembly (23)

H. Perceptual Speed	.17
B. Visualization	.11
Other	.06

Of all W-B tests, this one yielded the lowest communality, only about one-third of the variance of the test being identified. Probably a considerable proportion is specific, while estimates of its reliability suggest also a sizeable amount of error variance. In spite of these limitations, the Object Assembly test appears to be the best single test of the scale for the measurement of perceptual speed.

11. Digit Symbol (24)

C. Numerical Facility	.27
H. Perceptual Speed	.13
G. Fluency	.09
Other	.06

Another unexpected finding was the high loading on numerical facility for this test which does not involve any numerical operations. This finding is apparently in line with the views cited in the discussion of factor C. If what is called numerical facility is really the ability to manipulate numerical type symbols rapidly, whether for arithmetic operations or not, it would be of interest to know whether such a test could be used effectively to predict numerical aptitude regardless of cultural or language barriers.

Wechsler-Bellevue Factor Coverage. Considering only factors identified by this study, it appears that verbal comprehension, general reasoning, numerical facility, and education of conceptual relations are moderately well measured by individual subtests of the scale. Visualization, perceptual speed, mechanical knowledge, and fluency are measured less adequately, not appearing with high loadings in any subtest. These lists appear to reveal some tendency for the factor content to be more clearly identified in the verbal than in the performance area. On the other hand, however, both verbal and performance tests are poorly differentiated factorially, each factor appearing with significant loadings on several tests. At the same time, however, the factor coverage is probably adequate in depth only for verbal comprehension. This is not unexpected, however, since the scale was not constructed on a factor basis.

Summary and Conclusions

1. With view to factor analyzing the Wechsler-Bellevue scale in a matrix of reference variables, hypotheses were formulated, predicting, among other things, the appearance of at least six factors with significant loadings in subtests of the scale.

2. In line with the hypotheses, seven special reference tests were prepared and three Form II W-B tests were adapted for group administration.

3. These ten tests plus a science information test (included for local purposes) were administered to 356 students in grade 8A of a junior high school in Seattle. Age and Otis Beta scores were obtained for all students. The 11 subtests of the W-B scale were administered to 202 of the 356 students.

4. Intercorrelations of these 24 variables were factor analyzed by Thurstone's complete centroid method, yielding ten factors. Reference axes were rotated orthogonally by Zimmerman's graphic method.

5. Rotated factors were interpreted as: (a) verbal comprehension, (b) visualization, (c) numerical facility, (d) mechanical knowledge, (e) doublet (similarities), (f) general reasoning, (g) fluency, (h) perceptual speed, (j) education of conceptual relations, and (k) information.

6. Of the six factors hypothesized, one appeared as a leading factor in two W-B tests, four appeared as leading factors in one W-B test each, and one appeared only as secondary or tertiary in several tests.

7. All subtests of the W-B scale were found to be complex, Vocabulary, Arithmetic, and Picture arrangement being least so. Factorially, the Comprehension test proved to contribute little to the scale inasmuch as each factor appearing on it was better measured by some other subtest.

8. Inclusion of Form II of the Comprehension test resulted in the identification of a factor called fluency for which this test was the best single measure. This fact suggests that from the factor standpoint the Form II Comprehension test might

well supplant the Form I version in the scale. The fluency factor appears to have non-intellective characteristics such as interest Wechsler.

9. The analysis produced two other unexpected outcomes, (a) the Picture Arrangement test, although measuring reasoning, identified a reasoning factor quite distinct from that measured by arithmetic, and (b) the Digit Symbol test proved to be the best single measure in the scale for the numerical facility factor. This latter finding strengthened the belief that this factor is really one of ability to manipulate simple symbols rapidly.

REFERENCES

1. Balinsky, B. "An Analysis of the Mental Factors of Various Age Groups from Nine to Sixty", *Genetic Psychology Monograph*, XXIII (1941), 191–234.
2. Birren, J. E. "A Factor Analysis of the Wechsler-Bellevue Adult Intelligence Scale Given to an Elderly Population." *American Psychologist*, VI (1951), 398–399 (Abstract).
3. Birren, J. E. *A Factorial Analysis of the Wechsler-Bellevue Scale Given to an Elderly Population.* The Psychometric Laboratory, The University of Chicago, No. 73, 1951.
4. Burt, C. and John, Enid M. "Factorial Analysis of the Terman-Binet Tests, I." *British Journal of Educational Psychology*, XII (1942), 117–127.
5. Burt, C. and John, Enid M. "Factorial Analysis of the Terman-Binet Tests, II." *British Journal of Educational Psychology*, XII (1942), 156–161.
6. Cohen, J. "A Factor Analytic Comparison of Intelligence Test Performance of Different Neuropsychiatric Groups." *American Psychologist*, VI (1951), 334. (Abstract).
7. Comrey, A. L. "A Factorial Study of Achievement in West Point Courses." Educational and Psychological Measurement, IX (1949), 193–209.
8. Coombs, C. H. "A Factorial Study of Number Ability." *Psychometrika*, VI (1941), 161–189.
9. Davis, P. C. "A Factor Analysis of the Wechsler-Bellevue Scale, Form I, in a Matrix with Reference Variables." Unpublished doctoral dissertation, University of Southern California, 1952.
10. 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.

11. Dudek, F. J. "The Dependence of Factorial Composition of Aptitude Tests upon Population Differences among Pilot Trainees. II. The Factorial Composition of Test and Criterion Variables." Educational and Psychological Measurement, IX (1949), 95–105.
12. French, J. W. *The Description of Aptitude and Achievement Tests in Terms of Rotated Factors*. Chicago: University of Chicago Press, 1951.
13. Fruchter, B. "The Nature of Verbal Fluency." Educational and Psychological Measurement, VIII (1948), 33–47.
14. Garrett, H. E. "A Developmental Theory of Intelligence." *American Psychologist*, I (1946), 372–378.
15. Goldfarb, W. *An Investigation of Reaction Time in Older Adults.* New York: Teachers College, Columbia University Contributions to Education, No. 831, 1941.
16. Green, R. F. "A Factor Analytic Study of Reasoning Abilities." Unpublished doctoral dissertation, University of Southern California, 1951.
17. Guilford, J. P. *Psychometric Methods*. New York: McGraw-Hill, 1936.
18. Guilford, J. P., editor. *Printed Classification Tests*. Report No. 5. Army Air Forces Aviation Psychology Program Research Reports. Washington, D. C.: U. S. Government Printing Office, 1947.
19. Jones, L. V. "A Factor Analysis of the Stanford-Binet at Four Age Levels." *Psychometrika*, XIV (1949), 299–331.
20. Michael, W. B. "Factor Analyses of Tests and Criteria: a Comparative Study of Two AAF Pilot Populations." *Psychological Monographs, General and Applied*, LXIII (1949), No. 3.
21. Rabin, A. I. "The Use of the Wechsler-Bellevue Scales with Normal and Abnormal Persons." *Psychological Bulletin*, XLII (1945), 410–422.
22. Rabin, A. I. and Guertin, W. H. "Research with the Wechsler-Bellevue Test: 1945–1950." *Psychological Bulletin*, XLVIII (1951), 211–248.
23. Simkin, J. S. "An Investigation of Differences in Intellectual Factors between Normal and Schizophrenic Adults." *American Psychologist*, VI (1951), 335. (Abstract).
24. Taylor, C. W. "A Factorial Study of Fluency in Writing." *Psychometrika*, XII (1947), 239–262.
25. Thorndike, E. L. and Lorge, I. *The Teacher's Word Book of 30,000 Words*. New York: Bureau of Publications, Teachers College, 1944.
26. Thurstone, L. L. and Thurstone, Thelma G. *Factorial Studies in Intelligence*. Chicago: University of Chicago Press, 1941.

27. Thurstone, L. L. *Multiple-Factor Analysis*. Chicago: University of Chicago Press, 1947.
28. Wechsler, D. "Non-Intellective Factors in General Intelligence." *Psychological Bulletin*, XXXVII (1940), 444–445.
29. Wechsler, D. "Non-Intellective Factors in General Intelligence." *Journal of Abnormal and Social Psychology*, XXXVIII (1943), 101–103.
30. Wechsler, D. *The Measurement of Adult Intelligence*. Baltimore: Williams and Wilkins, 1944.
31. Wechsler, D. "Cognitive, Conative, and Non-Intellective Intelligence." *American Psychologist*, V (1950), 78–83.
32. Wright, Ruth E. "A Factor Analysis of the Original Stanford-Binet Scale." *Psychometrika*, IV (1939), 209–221.
33. Zimmerman, W. S. "A Simple Graphical Method for Orthogonal Rotation of Axes." *Psychometrika*, XI (1946), 51–55.

A STUDY OF THE EFFECTS OF "LIKINGNESS" AND LEVEL OF OBJECTIVITY ON PEER RATING RELIABILITIES[1]

GEORGE J. SUCI
University of Illinois

THEODORE R. VALLANCE
Human Research Unit No. 1, Fort Knox

AND

ALBERT S. GLICKMAN
U. S. Naval Personnel Research Field Activity, Washington

PERHAPS the most prominent use of peer ratings in a practical setting has been as a means of personnel assessment in military training situations. This is illustrated by the practice of having officer candidates rate one another as to various abilities or potentialities (1, 2, 3, 4).

Various writers have been concerned with the relation between peer ratings and "popularity" or "likingness", i.e., the extent to which peer ratings are influenced by personal attractiveness to the rater of the person being rated. That a significant relation may exist between likingness and peer rating has been shown in a study where the correlation between choice of roommate and choice of best all-around officer correlated .68 (4). No study, however, has concerned itself with the effect of the like-dislike variable upon the *reliability* of peer rating scores.

Problem

In the military setting, ideally, each peer group member should base his choices only upon attributes making solely for military leadership. However, it is possible that another

[1] This work was done by the Officer Personnel Research Program, at the U. S. Naval Schools Command, Newport, Rhode Island, under contract Nonr 890(00) between the American Institute for Research and the Office of Naval Research. The opinions expressed herein are those of the authors and do not necessarily reflect the opinions of the U. S. Navy.