Factor Analysis of the WAIS and Twenty French-Kit Reference Tests

Philip H. Ramsey
Hofstra University

Over a 3-year period, 20 reference tests were given to 114 undergraduate students ranging in age from 18 to 46. The WAIS was given to 107 of the students. Maximum likelihood factor analysis was performed on the 31-variable correlation matrix formed by the 11 WAIS subscales and 20 reference tests. A maximum likelihood test of significance supported the hypothesis of 10 factors. Significant split-sample factor reliabilities and WAIS subscale loadings were found on all 10 factors. The rotated factors in order of explained variance were Verbal Comprehension, Visualization, Memory Span, Syllogistic Reasoning, General Reasoning, Induction, Mechanical Knowledge, Number Facility, Spatial Orientation, and Associative Memory. The apparent contradictions between many previous analyses of the WAIS are discussed and a unified interpretation is presented.

Previous factor analyses of the Wechsler Adult Intelligence Scale (WAIS) seem to have presented widely differing underlying factor structures. The number of factors has ranged from 2 (Silverstein, 1969) to 10 or 11 (Saunders, 1959). Many of the previous factor analyses of the WAIS have been reviewed by Matarazzo (1972), who argued for three factors and severely criticized factor analytic research as highly subjective. All previous researchers seemed to have been able to produce factor analytic results which appeared to confirm their own theories of intelligence. Many of these theories seem to differ drastically. Wallbrown, Blaha, and Wherry (1974) have produced a three-factor solution including a *g* factor, while Saunders (1959) seems almost to have produced a separate factor for each of the 11 subscales of the WAIS. Such apparent discrepancies, coupled with "interpretations" of randomly generated data (Horn, 1967; Humphreys, Ilgen, McGrath, & Montanelli, 1969), seem to provide strong support for charges of subjectivity. It will be suggested later in the present work that the contradictions are more apparent than real.

Few of the previous factor analytic studies of the Wechsler tests have included reference tests to confirm or refute independently the many interpretations given to the various factors. An early study by Davis (1956) of the Wechsler-Bellevue was an exception, but it included only a limited number of reference tests. Cronbach (1970, p. 330) has suggested that at least 10 factors are measured to some degree by the Wechsler tests. The major purpose of the present investigation was to evaluate the contribution of each of the 10 specific factors mentioned by Cronbach by direct assessment, using reference tests produced by French, Ekstrom, and Price (1969).

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 2, No. 4 Fall 1978 pp. 505–517

Method

Subjects

Over a period of three years, data were collected on students in an undergraduate course in psychological testing at Hofstra University. In a five-semester period the number of students in each class was 22, 7, 41, 19, and 25, respectively. There were 44 men and 70 women, ranging in age from 18 to 46. Only 3 students were Black (2 males and 1 female). All 114 students were tested with 20 tests from the French Kit (French et al., 1969) for purposes of the course. All but 7 of the students also participated as subjects for further research by being tested with the WAIS. For the purpose of split-sample factor reliability estimation (Armstrong & Soelberg, 1968), the total sample was divided into two subsamples. Classes 1, 2, and 5 were combined to form Subsample A with 54 students; and Classes 3 and 4 were combined to form Subsample B, with 60 students. Marked differences in the classes should lead to systematic differences between Subsamples A and B, so comparable results from the two subsamples should be stronger support for factor reliability than results from two randomly divided subsamples.

Reference Tests

Twenty reference tests were used to establish 10 reference factors with 2 tests for each factor. They are as follows: Induction (I-1, I-2), Associative (Rote) Memory (Ma-2, Ma-3), Mechanical Knowledge (Mk-2, Mk-3), Memory Span (Ms-1, Ms-2), Number Facility (N-2, N-3), General Reasoning (R-2, R-4), Syllogistic Reasoning (Rs-1, Rs-3), Spatial Orientation (S-1, S-2), Verbal Comprehension (V-3, V-5), and Visualization (Vz-2, Vz-3). (For a more detailed description of the factors and tests see French et al., 1969.)

Procedure

The 20 reference tests were administered as group tests in six 1-hour sessions. All tests were administered according to the instructions and time limits specified in the test manual. The 107 students participating as subjects for WAIS testing were tested individually by graduate students who had just completed an intensive one-semester course in Wechsler testing. The 11 WAIS subscales were used in raw score form. Verbal IQ (VIQ), Performance IQ (PIQ), and Full Scale IQ (FIQ) were used in standard IQ units. The Arithmetic subscale was inadvertently omitted from the testing of the 19 students in Class 4. The VIQ and FIQ scores for those 19 students were calculated from the remaining subscales. All 19 students were in Subsample B, and the implications of these missing scores for the factor results are discussed. Subscale scores on one additional student were lost; but VIQ, PIQ, and FIQ scores for that student were based on all 11 subscales.

Three 31-variable correlation matrices were constructed, using the 20 reference tests and the 11 WAIS subscales. The matrices were for the total sample, Subsample A, and Subsample B.

The number of factors was determined by three conventional factor analytic procedures: (1) Kaiser's (1960) unit eigenvalue rule (based on the work of Guttman, 1954); (2) Jöreskog's (1963) maximum likelihood test of significance; and (3) Cattell's (1966) "scree" test of residual eigenvalues. The 10-factor unrestricted maximum likelihood solution from Jöreskog's program (UFABY3; Jöreskog & van Thillo, 1971) was rotated to fit a target matrix by the least squares procedure of Schönemann (1966). This procedure orthogonally rotates a given solution to any prespecified target solution. The target matrix was determined by placing 1's in the column corresponding to the appropriate factor for each reference test and O's in the other 9 columns. Tests 3 and 4 were targeted for 1's on the second factor, Tests 5 and 6 for the third factor, and so forth. Tests 1 and 2, the Induction tests, had a very low correlation (.254) which, although statistically significant, was not considered to be of large enough magnitude to adequately establish a common Induction factor.

Therefore, Tests 1 and 2 were treated as nonreference tests and not used to define a reference factor in the target matrix. On the other hand, two WAIS subscales—Digit Span and Vocabulary—were used as reference tests for Memory Span and Verbal Comprehension, respectively, since they were considered to represent adequate reference tests for those factors in their own right. The remaining 9 WAIS subscales, as well as Tests 1 and 2, were not targeted for any specific factor. Each of those 11 tests was targeted for the same loading on all 10 factors. Since orthogonal rotation leaves the communalities unchanged, the actual size of the equal target loadings makes little difference. Loadings of .30 were used, since that value gives a sum of squared loadings over the 10 factors of just under 1.0.

The two subsamples, A and B, were used to determine reliability of the factors. Due to the relatively small number of subjects, both subsample correlation matrices were singular; and maximum likelihood analyses could not be used. Each subsample was analyzed by principal factor analysis, using the largest off-diagonal correlation as the initial communality; and an iterative procedure was used to improve the communality estimate. Solutions were then rotated to least squares fits to the total sample solution, and corresponding factors from the two subsamples were compared, using both the salient variable similarity index (Cattell, Balcar, Horn, & Nesselroade, 1969) and the coefficient of congruence (Burt, 1948; Tucker, 1951).

A version of Dwyer's (1937) extension was used to estimate the loadings of five additional variables: VIQ, PIQ, FIQ, sex, and age. The exact procedure was the one described by Gorsuch (1974, pp. 129, 233).

In most factor analytic studies, the factors are identified by examining those variables with loadings above some minimum absolute value (usually .30) on a particular factor. That process was used in the present investigation but was supplemented by a more objective procedure. Jöreskog's (1970) confirmatory maximum likelihood analysis was used to test the fit of the final solution when "small" loadings were set to zero.

Results

Summary statistics for the 114 students on the 20 reference tests are presented in Table 1. Reliabilities agree with values reported in previous research (French, 1957; Fleishman & Dusek, 1971). The variance-covariance matrix of the 20-factor tests was compared for Subsamples A and B using the Box (1950) test, $\chi^2(210) = 205.53$, $p > .20$. The nonsignificance of the Box test indicates the absence of any marked difference between the two subsamples in the domain measured by these 20 tests, at least with regard to covariation. Of course, there could be differences on the WAIS measures not related to the reference tests.

Summary statistics on the WAIS measures, sex, and age are presented in Table 2. The range of WAIS FIQ scores was 101 to 137. As would be expected from a restricted range group, all reliabilities and standard deviations were somewhat lower than the values reported for the general population in the test manual (Wechsler, 1955). All reliabilities were statistically significant. Excluding the dichotomous sex variable, the remaining 35 variables were tested for skewness and kurtosis by testing the 3rd and 4th moments. The only variable to show significance in the departure from normality was age. Age showed both positive skewness and leptokurtic kurtosis.

The unit eigenvalue rule indicated 9 factors, while Cattell's scree test indicated 13 factors. Jöreskog's statistical test produced a significant chi square for 9 factors, $\chi^2(222) = 259.25$, $p = .044$. However, a nonsignificant chi-square was found for 10 factors, $\chi^2(200) = 226.02$, $p = .100$.

The total sample rotated maximum likelihood solution is presented in Table 3. All 10 factors were found to have significant split-sample reliability as indicated by the salient variable similarity index. Values of the coefficient of congruence were also quite high.

Table 1
Summary Statistics on All 20 Reference Tests for Total Sample (N = 114)

Test	Mean	S.D.	Range Min	Range Max	Reliability Split-Half	Reliability Alpha
1. I-1	20.22	4.00	6	29	.64	.70
2. I-2	11.07	4.42	-1	21	.70	.68
3. Ma-2	18.61	7.29	0	30	.81	.90
4. Ma-3	20.41	6.82	2	30	.88	.90
5. Mk-2	10.60	6.40	-3	27	.87	.79
6. Mk-3	13.54	6.01	3	30	.77	.75
7. Ms-1	12.54	3.34	5	20	---	.76
8. Ms-2	11.84	3.63	3	19	---	.80
9. N-2	37.64	13.21	7	70	.88	---
10. N-3	66.49	18.98	29	119	.92	---
11. R-2	13.33	4.73	3	25	.78	.78
12. R-4	17.90	4.18	6	29	.68	.68
13. Rs-1	8.20	7.89	-9	28	.75	.73
14. Rs-3	15.09	2.93	5	20	.53	.54
15. S-1	94.67	25.97	22	174	.87	---
16. S-2	15.18	9.16	-2	38	.73	.77
17. V-3	25.75	7.29	9	46	.83	.82
18. V-5	17.39	6.58	0	35	.86	.80
19. Vz-2	8.75	3.80	0	16	.71	.76
20. Vz-3	30.48	13.39	2	60	.85	.94

Note. Tests Ms-1 and Ms-2 have no subtests so split-half reliabilities were omitted. Alpha was omitted from the three highly speeded tests N-2, N-3, and S-1. All split-half reliabilities were from separately timed subtests.

The 10 factors presented in Table 3 can be easily identified, with the possible exception of Factor 1. The use of reference tests seems to have clearly established at least 9 of the 10 hypothesized factors. The use of Schönemann's least squares (or "Procrustes") rotation raises the question of how the present results compare with those of randomly generated data as reported by Horn (1967) and Humphreys et al. (1969). Interpolation from the randomly generated data of Humphreys et al. (1969) indicates that for 31 variables fitted to 10 factors with 2 tests per factor and an orthogonal rotation, the mean target loading from random data would be .397 with a standard deviation of .151. The present 20 target variables (Reference Tests 3 to 20 plus Digit Span and Vocabulary) had a mean loading of .677 with a standard deviation of .176. Comparing these mean loadings produced a significant difference, $t(38) = 5.26$, $p <$

Table 2
Summary Statistics for WAIS and Subject Variables

Variable	N	Mean	S.D.	R[a]
1. Information	106	22.29	3.17	.74
2. Comprehension	106	22.50	2.95	.56
3. Arithmetic	87	13.71	2.40	.61
4. Similarities	106	20.35	2.71	.37
5. Digit Span	106	13.62	2.24	.57
6. Vocabulary	106	62.82	7.79	.79
7. Digit Symbol	106	69.94	9.15	.83[b]
8. Picture Completion	106	15.90	2.51	.53
9. Block Design	106	37.66	7.21	.78
10. Picture Arrangement	106	26.47	4.52	.39
11. Object Assembly	106	34.24	6.60	.83
12. Verbal IQ (VIQ)	107	122.01	8.38	.84[c]
13. Performance IQ (PIQ)	107	112.53	10.06	.85[c]
14. Full Scale IQ (FIQ)	107	119.13	8.04	.88[c]
15. Sex	114	.39	.49	---
16. Age	107	22.94	5.34	---

Note. Raw scores were used on all WAIS subscales.
[a]Reliabilities for all WAIS subscales except Digit Symbol were calculated from Flanagan's (Kelley, 1942) formula for split-half reliability with unequal halves.
[b]Digit Symbol reliability was estimated using the restricted range formula employed by Wechsler (1955, p. 12) for the standardization sample.
[c]Reliabilities for all IQ scores were estimated by Nunnally's (1967, p. 229) formula 7-10 for the reliability of a composite score.

.001. Clearly, the present results showed much higher loadings than those produced by random data.

Most factor analytic studies, as mentioned earlier, set some minimum absolute loading size, usually .30, as the smallest absolute loading a variable may have and still be of use in the interpretation of the factor. Only variables with absolute loadings above the minimum are used to identify each factor. In the present study, 9 of the 10 predicted factors were easily identified by this traditional method. Factor 1 was identified as Induction on the basis of the Test 2 (I-2) loading of .34, but the near zero loading of .05 for Test 1 (I-1) suggests that the identification must be considered tentative.

Once the factors are identified, the process is generally reversed and the "factorial content" of each variable is evaluated by examining all factors on which a particular variable has absolute loadings above the minimum .30. The computer program ACOVS (Jöreskog, Gruvaeus, & van Thillo, 1971) was used to provide a more objective minimum value for the contribution of a

Table 3
Rotated Maximum Likelihood Factor Matrix for Total Sample (N = 114) with 31 Variables

Variable	Factor 1 I	2 Ma	3 Mk	4 Ms	5 N	6 R	7 Rs	8 S	9 V	10 Vz	Communality h^2
1. I-1	05	10	-13	23	38[b]	33[b]	06	14	-02	17	387
2. I-2	34[b]	24	-01	13	-03	38[b]	06	26	-05	09	419
3. Ma-2	-06	81[a]	-07	10	-02	09	-06	08	01	-13	715
4. Ma-3	-06	73[a]	-13	12	01	12	20	-02	11	-08	642
5. Mk-2	-15	-06	78[a]	03	-06	12	00	21	26	14	784
6. Mk-3	00	-18	78[a]	07	-09	26	05	08	08	07	739
7. Ms-1	-01	05	00	82[a]	-04	04	-03	06	12	-12	716
8. Ms-2	00	10	01	77[a]	12	12	07	01	05	03	647
9. N-2	00	-09	-02	27	72[a]	32[b]	01	12	19	-11	769
10. N-3	-10	09	-07	13	80[a]	28	-11	-01	22	-23	861
11. R-2	-09	07	17	15	26	85[a]	28	-01	10	21	995
12. R-4	-10	06	17	13	24	48[a]	09	30[b]	21	09	505
13. Rs-1	09	05	10	27	-08	35[b]	41[a]	07	23	01	449
14. Rs-3	-11	19	09	06	05	19	38[a]	-08	20	11	304
15. S-1	-01	02	17	20	08	13	-06	36[a]	-07	34[b]	351
16. S-2	-04	09	23	08	07	30[b]	02	42[a]	19	32[b]	485
17. V-3	-09	04	13	07	19	11	16	06	83[a]	-03	807
18. V-5	-15	01	09	09	07	19	02	-10	87[a]	-09	859
19. Vz-2	12	11	16	04	-10	23	04	35[b]	11	53[a]	531
20. Vz-3	-05	-14	28	13	-07	34[b]	17	44[b]	00	56[a]	771
21. Information	30[b]	-04	22	05	21	31[b]	06	24	45[b]	09	555
22. Comprehension	16	08	11	-02	15	-26	62[b]	14	33[b]	-02	651
23. Arithmetic	28	04	02	23	27	54[b]	24	28	11	-09	656

(continued on next page)

Table 3 (continued)

24. Similarities	34[b]	14	17	-02	-09	-14	46[b]	20	20	-11	494
25. Digit Span	-19	00	04	79[a]	22	12	11	06	09	-02	755
26. Vocabulary	09	03	12	13	09	05	11	05	85[a]	00	787
27. Digit Symbol	-12	35[b]	00	10	30[b]	-03	-10	09	-19	14	314
28. Picture Completion	01	23	49[b]	12	-01	-17	19	19	18	33[b]	546
29. Block Design	42[b]	17	30[b]	22	20	29	-04	-09	17	70[b]	994
30. Picture Arrangement	15	-02	03	07	12	00	12	19	15	33[b]	223
31. Object Assembly	-02	20	17	12	09	00	18	07	06	50[b]	389
Per Cent of Variance Explained, 11 Subscales	516	250	423	718	324	551	705	295	1107	915	5786
Per Cent of Variance Explained, 31 Variables	262	544	649	787	610	837	422	385	985	663	6161
Coefficient of Congruence	88	76	77	86	77	91	79	61	89	80	
s-Index	42**	38*	52**	62**	53**	70**	44**	50**	48**	56**	
Dwyer's Extension											
32. VIQ	22	09	16	40	20	13	47	22	48	-02	803
33. PIQ	15	27	31	21	27	04	09	15	15	60	722
34. FIQ	22	21	27	38	28	10	34	22	39	33	815
35. Sex	-05	-20	47	16	-03	26	-07	08	08	-01	373
36. Age	09	-02	05	04	15	-07	-03	-02	34	-14	181

Note. All decimals are omitted. All numbers have two decimal places except communalities which have three. All per cents (whether 3 or 4 digits) have two decimal places.

[a]These positions had target values of 1. All other positions in the same row had target values of 0. All positions in other rows had target values of .30 for all ten factors.

[b]These loadings are .30 or above even though targeted for only .30 or 0.

*$p < .05$.

**$p < .01$.

variable to a factor (or vice-versa). As expected, the full solution in Table 3—after all loadings below .10 were set to 0—produced a good fit of the factor solution to the original correlation matrix, $\chi^2(379) = 354.69$, $p = .810$. However, if all loadings less than .13 were set to 0, the departure of the fit just missed significance at the .05 level, $\chi^2 = 422.48$, $p = .061$. Setting all loadings less than .15 to 0 produced a highly significant departure, $\chi^2(379) = 463.72$, $p = .002$. To avoid a statistically significant departure, loadings of .13 or greater cannot be ignored.

This seems to indicate that consideration only of loadings of .30 or greater ignores a great deal of important information in the factor solution. It should be noted that these small loadings (.13 to .30) are not of much importance in the identification of factors, since they represent so little of the variance of the factor. On the other hand, they may be very important in identifying the factorial content of a variable. The shared variance of a variable might be spread thinly over many factors. To adequately reproduce that variable, one would need to include information from factors with loadings even as low as .13. Unfortunately, none of the specific loadings (in the range from .13 to .30) can be accepted with much confidence. Although the aggregate effect of all these loadings may be crucial, any specific loading could be a chance value.

The large number of loadings in Table 3 with absolute values of .13 and above suggests considerable complexity in the factor structure for most of the WAIS subscales. Most of the subscale communalities are reasonable except for Block Design (.994), which is clearly inflated. Some of the other communalities, such as Digit Symbol (.314) and Picture Arrangement (.223), seem too low; and these subscales would probably load on factors other than the present 10. Estimated loadings for VIQ, PIQ, FIQ, sex, and age from Dwyer's extension are also presented in Table 3. A number of these exceed .30, and quite a few more exceed .13 in absolute value.

Discussion

Although the 3 empirical procedures ranged from 9 to 13 in their indication of the number of factors, they seem to give general support to the use of 10 factors. In particular, the maximum likelihood solution showed a nonsignificant residual matrix after the extraction of 10 factors. Additional reference tests could probably be found to establish 2 or 3 more factors. However, the present set of 31 variables seemed to be quite adequately represented by a 10-factor solution.

Factor 9, Verbal Comprehension, is the strongest single factor in the rotated solution presented in Table 4 and explained 11.07% of the variance of the 11 WAIS subscales. Verbal Comprehension was also identified by Davis (1956) in his analysis of the Wechsler-Bellevue. He reported similar loadings on several subscales—Vocabulary (.601), Information (.423), and Similarities (.310).

Factor 10, Visualization, was the second strongest factor, explaining 9.15% of the WAIS subscale variance. Davis (1956) also identified a Visualization factor with high loadings on Block Design (.439), Picture Completion (.384), and Object Assembly (.338). Even Saunders's (1959) analysis showed Block Design and Object Assembly to appear on the same factor, as they do on Factor 10 in the present case. An interesting feature of the present Visualization factor is that the original targeted reference tests, Vz-2 and Vz-3, overlapped in factor structure with Spatial Orientation (Factor 8), while Block Design and Object Assembly did not. These WAIS subscales may be better discriminators of the Visualization and Spatial Orientation factors than are the reference tests.

The third strongest factor was Factor 4, Memory Span, explaining 7.18% of the subscale variance. Berger, Bernstein, Klein, Cohen, and Lucas (1964) and Cohen (1957) have reported a similar factor. None of the previous analyses seem to have shown quite as distinct a Memory Span factor as the present results. On the other hand, none of the previous analyses (not even

Davis, 1956) included reference tests for Memory Span. It would appear that the factor can be established if it is specified as in a target matrix. Some writers, such as Matarazzo (1972, pp. 273–274), have argued for the inclusion of such a factor in factor analytic results but do not identify it as Memory Span.

Factor 7, Syllogistic Reasoning, was the fourth strongest factor, explaining 7.05% of the WAIS subscale variance. This factor and the remaining ones have not been widely reported in previous analyses of the WAIS. Davis (1956) did not include reference tests for this factor; the closest factor he reported was a Similarities factor which had large loadings only from two similarities tests. In other analyses (Cohen, 1957; Berger et al., 1964) this factor seems to have been combined with Verbal Comprehension to form a single verbal factor. The identification of Syllogistic Reasoning as the dominant factor in Similarities and Comprehension does not seem to have been determined previously.

Factor 6, General Reasoning, was the fifth strongest factor, explaining 5.51% of the WAIS subscale variance. It resembles the fourth factor reported by Berger et al. (1964) in some of the analyses they reported, but in others it seems to have been combined with Memory Span. Davis (1956) identified the same factor in his Wechsler-Bellevue analysis, finding a high loading for Arithmetic (.571). However, Davis also reported a moderately high loading for Comprehension (.331), which had a slightly negative loading here of −.26.

Factor 1, Induction, was the sixth strongest factor, explaining 5.16% of the WAIS subscale variance. Its identification as Induction may be questionable, since the two reference tests—I-1 and I-2—failed to establish a clear factor. The heavy loading for Block Design (.42) is questionable because of its clearly inflated communality of .994. To identify Factor 1, five tests with moderately high loadings (.28 or higher) can be examined, as shown in Table 3—I-2 (.34), Information (.30), Arithmetic (.28), Similarities (.34), and Block Design (.42). Davis (1956) reported a similar factor with loadings on three subscales: Information (.560), Arithmetic (.318), and Block Design (.305). Davis identified his factor as Information after the highest subscale loading. For lack of a better name, Induction was retained on the basis of the I-2 loading.

Factor 3, Mechanical Knowledge, explained 4.23% of the WAIS subscale variance. Davis (1956) again identified the same factor, although he found high loadings only for Block Design (.411), Arithmetic (.340), and Digit Span (.336) among the Wechsler-Bellevue subscales.

Factor 5, Number Facility, explained 3.24% of the WAIS subscale variance. Davis (1956) identified the same factor with high loadings for Digit Symbol (.524) and Arithmetic (.359). Davis also found a high loading for age (.320). The extension analysis would appear to agree, with an age loading of .15; but Davis reversed his scoring of age to give younger subjects the higher score. The disagreement in loadings between Davis and the present results is probably due to differences in the age groups examined. Davis used 12- to 16-year-old junior high school students. The students in the present sample ranged in age from 18 to 46, with the majority of them in their 20's.

Factor 8, Spatial Orientation, was the ninth factor in size and explained only 2.95% of the WAIS subscale variance. Davis (1956) did not report such a factor, but he acknowledged the absence of any reference tests for memory or spatial relations. None of the WAIS subscales showed loadings above .30 in Table 3. However, a number of them exceeded .13: Information (.24), Comprehension (.14), Arithmetic (.28), Similarities (.20), Picture Completion (.19), and Picture Arrangement (.19). Spatial Orientation seems to play a statistically significant, although relatively minor, role in a number of the WAIS subscales.

Factor 2, Associative (Rote) Memory, was the weakest of the 10 factors, explaining only 2.50% of the WAIS subscale variance. It was not identified by Davis (1956) nor by any other previous factor analytic study which has been reviewed here. Digit Symbol (.35) has the most prominent loading.

The only factor reported by Davis (1956) which represents a likely candidate for an eleventh factor is his Perceptual Speed. He reports high loadings for Object Assembly (.415), Block Design (.385), and Digit Symbol (.366). Object Assembly and Digit Symbol both have low communalities in the present results, with values of .389 and .314, respectively. Inclusion of a perceptual speed reference test might have established such a factor.

Five of the present factors—Verbal Comprehension, Visualization, General Reasoning, Number Facility, and Mechanical Knowledge—seem to represent clear replication of the results of Davis (1956). Replication is particularly important because of the many differences between the present work and that of Davis. Three more factors—Memory Span, Syllogistic Reasoning, and Induction—are similar to factors found by Davis or others (Cohen, 1957; Berger et al., 1964). The two remaining factors, Spatial Orientation and Associative Memory, have not previously been shown to contribute to WAIS measures. There seems to be good support for Cronbach's specification of the 10 factors which were investigated. One or two additional factors may be justified, with Perceptual Speed being the most likely candidate.

The extension analysis provided some important additional information about the WAIS factor structure. For example, age had a major factor loading only on Verbal Comprehension (.34). This loading seems reasonable, since this ability is likely to increase with education and experience.

Present correlations of sex with VIQ, PIQ, and FIQ were all non-significant with respective values of .082, .114, and .116. This is the expected result, since the WAIS was constructed in such a manner as to eliminate sex differences from IQ scores (Matarazzo, 1972, p. 352). However, sex differences have been noted in the subscales. The combined 1955 standardization data on 850 males and 850 females aged 16 to 64 showed small, but statistically significant, differences on 8 of the 11 subscales (Matarazzo, 1972, p. 355). Converting those mean difference critical ratios to correlations, in order to facilitate comparison to present results, produces quite small correlations. Males were better on Information (.090), Arithmetic (.174), Comprehension (.053), Picture Completion (.119), and Block Design (.067). Females were better on Similarities (−.054), Vocabulary (−.059), and Digit Symbol (−.177). None of these correlations would have exceeded the .05 critical value, .189, required for significance in the much smaller total sample size of 106 in the present study. The .189 critical value was exceeded by only 4 of the present 11 correlations between sex and subscale scores. As in the standardization data, males were higher on Information (.311), Arithmetic (.192), Picture Completion (.250), and Block Design (.194). Females did not score significantly higher on any of the subscales.

Higher male scores on Arithmetic were reflected in the moderately high loading of sex on General Reasoning (.26). Males also scored above females in Mechanical Knowledge, as shown by the loading of .47. Smaller differences were present in the memory factors, with higher female scores on Associative Memory (−.20) and higher male scores on Memory Span (.16).

The various IQ measures in Table 3 showed surprisingly equal loadings on all factors. The heavy loading of PIQ on Visualization (.60) is to be expected, since so many performance subscales load on that factor. The most unexpected loading is probably the low value of FIQ on General Reasoning (.10), which occurred despite the rather high loading of Arithmetic on General Reasoning (.54). This low loading cannot be easily explained as attributable to the 19 students with missing Arithmetic scores. An extension analysis based on the 87 students with no missing scores produces almost identical results, with FIQ loading .11 on General Reasoning. However, a few such unexpected results could be due to chance and require replication.

The present investigation has attempted to detect substantial loadings for WAIS measures on the 10 relatively independent factors. The attempt has been generally successful; however, in reviewing previous factor analytic results, it was noted that most investigators have been equally able to produce results in agreement with their own theories. The most likely explanation for the success of so many different theories must be that all of the theories are compatible with the data. This would suggest that from the point of view of reproducing the original correlation matrix, any of the factor solutions would be adequate. The accuracy of the reproduction is probably determined primarily by the number of factors.

An unrotated principal components solution for the present data offers some useful information about the number of factors. The first unrotated component had an FIQ loading of .883 and was a good estimate of the *g* factor. It explained 24% of the variance of the 11 WAIS subscales. Any investigator whose interests can be served by a single measure of general intelligence could use a one-factor solution with FIQ as the measure of that factor.

Investigators who desire a two-factor solution with a verbal and performance factor could use the results of Silverstein (1969) and measure the two factors by VIQ and PIQ. The present unrotated components solution showed 31% of the WAIS subscale variance being explained by the first two components.

Three- and four-factor solutions have also been presented (Berger et al., 1964; Cohen, 1957; Wallbrown et al., 1974). The present unrotated third and fourth components show cumulative explained variances of 42% and 50%. It is not immediately obvious how to best measure those factors, but empirical estimates could be constructed.

At least three 10-factor solutions are available, the present solution and those of Davis (1956) and Saunders (1959). The 10 unrotated components of the present analysis explained 76% of the WAIS subscale variance, with no less than 3% of the variance being explained by any one component. One argument for the use of 10 or more factors is that the reliability of the subscales seems to be sufficiently large to justify explaining more than 70% of the variance. However, it must be kept in mind that both the number and structure of factor solutions are influenced by the number and type of tests analyzed. Additional or different reference tests might change many of the present observations. Despite these limitations, there seems to be good justification for a 10-factor solution.

The present results have already been shown to agree closely with those of Davis (1956). For a 10-factor solution of the WAIS, an investigator could choose either the Saunders (1959) type factor solution or the reference test solution used here and by Davis. The advantage of the reference test solution is that the WAIS factors can then be related to a larger body of factor analytic research (French, 1957; Fleishman & Dusek, 1971).

It is not clear from the present analysis how many, if any, of the 10 factors can be adequately measured by WAIS subscales. Vocabulary and Digit Span must be good candidates for measuring Verbal Comprehension and Memory Span. Visualization and General Reasoning also seem likely as abilities measurable by WAIS subscales. The reliabilities of any such measures would need to be determined. Perhaps an even more important consideration is the one raised by Silverstein (1969) regarding the predictive utility of any such factor measure.

Several important issues must be considered in the final evaluation of the present investigation. Many previous researchers (Cohen, 1957; Silverstein, 1969; Wallbrown et al., 1974) have managed to maintain 18 or more subjects per variable by restricting their analyses to the standardization data for the WAIS. Others (Davis, 1956; Saunders, 1959) have typically had 8 to 12 subjects per variable. Unfortunately, *no* reference tests have been included in any previous factor analyses of the WAIS (Davis only included a few reference tests, and he analyzed the Wechsler-Bellevue). All previous analyses were also highly subjective both in the rotation and identification of the factors. The present study had only 3.5 subjects per variable; but most of the variables were reference tests, the factors were defined before the data were collected, and the rotations were completely objective. Actually, the domain being analyzed was only the 11 subscales of the WAIS for which there were from 8 to 11 subjects for every variable.

There are some difficulties due to missing data in the present study. However, previous studies (e.g., Davis, 1956) have had the same problem when a large number of variables were being used; and in spite of these difficulties, the present results agree well with previous results, where comparisons are possible. In fact, the success of the present analysis in producing reasonable results provides some limited support to the practice of calculating WAIS IQs when the score on a single subscale is missing.

The use of graduate students as examiners might also be challenged, even though these students were carefully trained and closely supervised. Again, previous work suggests there should be no problem, since Saunders (1959) reported that half his data were collected by an initially inexperienced examiner; and when the halves were analyzed separately, Saunders found no important difference.

The present results should be of considerable use to applied psychologists, despite the outdating of the WAIS by the soon-to-be-published revision. The present version of the WAIS will almost certainly be used for many years, since that has happened with previous tests. Any differences or agreements between the old and new WAIS in factor structure will be most clearly shown by factor analyses based on reference tests.

References

Armstrong, J. S., & Soelberg, P. On the interpretation of factor analysis. *Psychological Bulletin,* 1968, *70,* 361–364.

Berger, L., Bernstein, A., Cohen, J., & Lucas, G. Effects of aging and pathology on the factorial structure of intelligence. *Journal of Consulting Psychology,* 1964, *28,* 199–207.

Box, G. E. P. Problems in the analysis of growth and wear curves. *Biometrics,* 1950, *6,* 362–389.

Burt, C. L. The factorial study of temperamental traits. *British Journal of Psychology* (Statistical Section), 1948, *1,* 178–203.

Cattell, R. B. The meaning and strategic use of factor analysis. In R. B. Cattell (Ed.), *Handbook of multivariate experimental psychology.* Chicago: Rand McNally, 1966.

Cattell, R. B., Balcar, K. R., Horn, J. L., & Nesselroade, J. R. Factor matching procedures: An improvement of the *s* index; with tables. *Educational and Psychological Measurement,* 1969, *29,* 781–792.

Cohen, J. The factorial structure of the WAIS between early adulthood and old age. *Journal of Consulting Psychology,* 1957, *21,* 283–290.

Cronbach, L. J. *Essentials of psychological testing* (3rd ed.). New York: Harper & Row, 1970.

Davis, P. C. A factor analysis of the Wechsler-Bellevue Scale. *Educational and Psychological Measurement,* 1956, *16,* 127–146.

Dwyer, P. S. The determination of the factor loadings of a given test from the known factor loadings of other tests. *Psychometrika,* 1937, *2,* 173–178.

Fleishman, J. J., & Dusek, E. R. Reliability and learning factors associated with cognitive tests. *Psychological Reports,* 1971, *29,* 523–530.

French, J. W. The factorial invariance of pure-factor tests. *Journal of Educational Psychology,* 1957, *48,* 93–109.

French, J. W., Ekstrom, R. B., & Price, L. A. *Manual for kit of reference tests for cognitive factors* (Rev. ed.). Princeton, NJ: Educational Testing Service, 1969.

Gorsuch, R. L. *Factor analysis.* Philadelphia: W. B. Saunders Co., 1974.

Horn, J. L. On subjectivity in factor analysis. *Educational and Psychological Measurement,* 1967, *27,* 811–820.

Humphreys, L. G., Ilgen, D., McGrath, D., & Montanelli, R. Capitalization on chance in rotation of factors. *Educational and Psychological Measurement,* 1969, *29,* 259–271.

Jöreskog, K. G. *Statistical estimation in factor analysis.* Stockholm: Almqvist & Wiksell, 1963.

Jöreskog, K. G. A general method for analysis of covariance structures with applications. Part I: General methodology; Part II: Applications. *Biometrika,* 1970, *57,* 239–251.

Jöreskog, K. G., & van Thillo, M. *New algorithms for factor analysis by unweighted least squares, generalized least squares and maximum likelihood* (RM-71-5). Princeton, NJ: Educational Testing Service, 1971.

Jöreskog, K. G., Gruvaeus, G. T., & van Thillo, M. *ACOVS, A general computer program for analysis of covariance structures* (RB-70-15). Princeton, NJ: Educational Testing Service, 1971.

Kaiser, H. F. The application of electronic computers to factor analysis. *Educational and Psychological Measurement,* 1960, *20,* 141–151.

Matarazzo, J. D. *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore, MD: Williams & Wilkins, 1972.

Saunders, D. R. On the dimensionality of the WAIS battery for two groups of normal males. *Psychological Reports,* 1959, *5,* 529–541.

Schönemann, P. H. A generalized solution of the orthogonal Procrustes problem. *Psychometrika,* 1966, *31,* 1–10.

Silverstein, A. B. An alternative factor analytic solution for Wechsler's intelligence scales. *Educational and Psychological Measurement,* 1969, *29,* 763–767.

Tucker, L. R. *A method for synthesis of factor analysis studies.* (Personnel Research Section Report No. 984). Washington, DC: Department of the Army, 1951.

Wallbrown, F. H., Blaha, J., & Wherry, R. J. The hierarchical factor structure of the Wechsler Adult Intelligence Scale. *British Journal of Educational Psychology,* 1974, *44,* 47–56.

Wechsler, D. *Manual for the Wechsler Adult Intelligence Scale.* New York: Psychological Corporation, 1955.

Acknowledgments

The author thanks Dr. Howard Kassinove and Dr. Marie Meier for assisting in the supervision of the WAIS testing.

Author's Address

Philip H. Ramsey, Psychology Department, Hofstra University, Hempstead, NY 11550 (After September 1978: Psychology Department, Queens College of the City University of New York, Flushing, NY 11367).