

WISC-R Subscale Patterns of Abilities of Blacks and Whites Matched on Full Scale IQ

Cecil R. Reynolds
Texas A & M University

Arthur R. Jensen
University of California, Berkeley

Groups of 270 black and 270 white children drawn from the national stratified random sample used in the standardization of the Wechsler Intelligence Scale for Children—Revised (WISC-R) were matched on age, sex, and WISC-R Full-Scale IQ to facilitate investigation of the patterns of specific cognitive abilities, as measured by the 12 subtests of the WISC-R, between the two racial groups. Multivariate analysis of the patterns of subtest differences between whites and blacks and group comparisons on three orthogonalized factor scores (verbal, performance, memory) show small but reliable average white-black differences in patterns of ability. The IQ-matched racial groups show no significant difference on the verbal factor; whites exceed blacks on the performance (largely spatial visualization) factor; blacks exceed whites on the memory factor.

At least since the seminal study by Lesser, Fifer, and Clark (1965), differential psychologists have been aware that various racial or ethnic groups differ from one another, on average, more on some mental tests than on others. A battery of various tests thus shows different mean profiles or patterns of the measured abilities for different groups. Lesser, Fifer, and Clark (1965) administered tests of verbal, reasoning, number, and spatial abilities to 6-8-year-old Chinese, Jewish, black, and Puerto Rican children in New York City. The four groups showed distinctly different patterns of ability. The most striking finding of the study is that groups of high and low socioeconomic status (SES) *within* each ethnic group showed almost identical patterns of ability. SES in this study is related to overall *level* of ability rather than to differential profiles of abilities, which are related to ethnicity.

A recent review (Willerman, 1979) of the major literature on this topic cites seven studies. In a more recent critique, Jensen (1980, pp. 729-736) has elaborated a number of the inherent methodological problems with such studies of differential patterns of abilities among various populations, making

the psychological and psychometric interpretation of such differences highly ambiguous. The most serious ambiguity probability results from the fact that the groups may actually differ on only one or a very few *independent factors* of ability, and because the various tests in the battery have different loadings on these few factors, it could cause the groups to differ from one another to varying degrees on each of a large number of tests. If two groups differed *only* in Spearman's *g* (the general intelligence factor), but differed in no other ability factors, and if the groups were compared on a dozen or so tests which differed markedly in their *g* loadings, the groups would show distinctly different profiles of scores on the various tests. They would also show different profiles even if all the tests had identical *g* loadings but differed markedly in reliability; the tests' reliabilities would be directly related to the magnitudes of the group differences when these are expressed in standard score units.

Several studies (Jensen, 1980, pp. 536-552, 732-736; Note 1) have shown that the magnitude of the average difference between blacks and whites on various tests is substantially related to the tests' *g* loadings (i.e., first principal component or first principal factor), which accords with Spearman's (1927, pp. 379-380) conjecture that the black-white difference in tests of mental ability reflects mainly a difference in *g* rather than in any of the narrower group factors

Requests for reprints should be sent to Cecil R. Reynolds, Department of Educational Psychology, Texas A & M University, College Station, Texas 77843 or to Arthur R. Jensen, Institute of Human Learning, University of California, Berkeley, California 94720.

measured by the tests or the tests' specificity. The existing evidence seems to bear this out in the main, and Jensen (1980) has suggested that "With our present evidence and the lack of any proper profile studies, . . . it would be difficult to make a compelling argument that blacks and whites differ on any abilities *other* than *g* in both its fluid and crystallized aspects" (p. 732).

Do blacks and whites differ in any abilities other than *g*? Of course, we are here dealing only with phenotypic differences. We are investigating the existence and nature of the phenotypic ability differences between whites and blacks, whatever their causes, and asking specifically whether there are population differences in abilities other than *g*. The question is of interest for practical as well as theoretical reasons. If there are true differences in patterns of ability, it could mean that the total score derived from a composite of a number of different subtests, as in the Wechsler intelligence scales, is not composed of equal parts of the same factors for blacks and whites. Hence, blacks and whites with exactly the same Full Scale IQ on a Wechsler scale may obtain their scores in characteristically quite different ways. Consequently, somewhat different interpretations or predictive inferences might be warranted for two individuals with the same Full Scale IQ but different patterns of ability.

In a recent study of white-black differences in subscale patterns on the Wechsler Intelligence Scale for Children—Revised (WISC-R; Wechsler, 1974), Vance, Hankins, and McGee (1979) reported that blacks earn their highest level of performance on the verbal subtests, a finding in sharp contrast to the popular belief that blacks are relatively disadvantaged on verbal as contrasted to nonverbal tests. Their finding, however, accords with many other studies of the verbal-nonverbal test score differences among blacks as compared with white. (For a comprehensive review of these studies, see Jensen, 1980, pp. 527–533.) These studies, however, have either failed to use representative samples of the black and white populations in the United States or to take account of the relative *g* loadings of verbal and nonverbal tests. It could well be that nonverbal tests are usually more *g* loaded than

verbal tests, thereby showing larger average white-black differences, in accord with Spearman's hypothesis.

The present study seeks to refine our knowledge of black-white differences in patterns of ability on the WISC-R by examining subtest differences between a large national, stratified (to match the 1970 U.S. Census) random sample of black children and a comparably sized group of white children selected from a national, stratified random sample so as to match the distribution of Full Scale IQs of the black sample as closely as possible. The Full Scale IQ is a quite close, although not perfect, estimate of the general factor of the WISC-R. Significant black-white differences on the various WISC-R subtests hence cannot be attributed to differences in general level of ability, on which the groups are almost perfectly matched but would reflect population differences in factors specific to each test or to group factors common to certain groups or types of subtests. The latter possibility is examined by comparing the two racial samples on three orthogonal factor scores derived from the 12 WISC-R subscales. Thus, the present study corrects many of the methodological and interpretive pitfalls of previous studies of cross-racial ability patterns.

Method

Subjects

The WISC-R standardization sample of 2,200 children between the ages of 6 and 16 ½ years provided the children for the study. These children were chosen in a stratified, random sampling procedure to be representative of the United States population at large, based on 1970 census figures. The sample was stratified on the basis of age, sex, race, SES, geographic region of residence in the U.S., and urban versus rural residence. The sample contained 305 blacks. This sample is described in great detail elsewhere (Kaufman & Doppelt, 1976; Reynolds & Gutkin, 1979; Wechsler, 1974). To obtain the sample for use in the present study, an attempt was made to match each of the 305 black children with a white child on the basis of age (within 1 year, even though age is uncorrelated with scaled scores on the WISC-R; see Reynolds & Gutkin, 1979), sex, and Full Scale IQ (within 1 standard error of measurement, about 3 IQ points). Using this matching procedure, 270 exact matches were obtained. In the case of multiple matching whites for any black child, the children were matched on the basis of SES as determined by the father's occupation, a matching condition invoked only

rarely. Since random samples of whites and blacks differ significantly on all subtests and scales of the WISC-R (Reynolds & Gutkin, 1981), the matching procedure provides a more accurate, overall level-free, picture of the differences in *pattern* of performance between whites and blacks.

Procedure

Each of the WISC-R subtests, in addition to measuring a general factor of ability common to all of the subtests, also reliably measures certain more distinct abilities—broad group factors and narrower abilities that are specific to each subtest (Kaufman, 1975, 1979). Examination of white-black differences in the 12 subtests, after matching white and black subjects on chronological age and Full Scale IQ, was based on a multivariate analysis of variance of the group differences simultaneously over all 12 subtests and the Verbal, Performance, and Full Scale IQs, followed by significance tests of the groups' mean differences on each of the scores. Also, significance tests were done on the group mean differences on each of three uncorrelated factor scores representing the main group factors that contribute to WISC-R variance: verbal, performance, and memory.

Results and Discussion

Subtest Differences

Table 1 shows the means and standard deviations of the scaled scores (for the entire

national standardization sample, $M = 10$, $\sigma = 3$) of the matched white and black groups (each with $n = 270$) on each of the WISC-R subtests, the uncorrected mean group difference ($D = \text{white } M - \text{black } M$), and the univariate F tests of the significance of the differences. The Verbal, Performance, and Full Scale IQs are also shown to indicate the degree to which the matching procedure affects these scores. The mean Full Scale IQs of the matched whites and blacks differ only $.03\sigma$. Unlike the majority of other studies in which the black and white groups are not matched in general level of ability, these groups matched on FS IQ show a negligible mean difference ($D = -.02$, which is $-.002\sigma$) in Verbal IQ, further substantiating the effectiveness of the matching procedure in removing general ability from consideration.

Because the 15 variables in Table 1 are highly intercorrelated, univariate comparisons of the groups on the separate subtests depend first on establishing the overall significance of the white-black differences among the 15 pairs of means. A multivariate analysis of variance reveals that the patterns of subtest means differ significantly between whites and blacks, $F(15, 524) =$

Table 1
Means, Standard Deviations, and Univariate F s for Comparison of Performance on Specific WISC-R Subtests by Groups of Blacks and Whites Matched on WISC-R Full Scale IQ

WISC-R variable	Whites		Blacks		D^a	F^b
	M	SD	M	SD		
Information	8.24	2.62	8.40	2.53	-.16	.54
Similarities	8.13	2.78	8.24	2.78	-.11	.22
Arithmetic	8.62	2.58	8.98	2.62	-.36	2.52*
Vocabulary	8.27	2.58	8.21	2.61	+.06	.06
Comprehension	8.58	2.47	8.14	2.40	+.44	4.27**
Digit Span	8.89	2.83	9.51	3.09	-.62	6.03***
Picture Completion	8.60	2.58	8.49	2.88	+.11	.18
Picture Arrangement	8.79	2.89	8.45	2.92	+.34	1.78*
Block Design	8.33	2.76	8.06	2.54	+.27	1.36
Object Assembly	8.68	2.70	8.17	2.90	+.51	4.41**
Coding	8.65	2.80	9.14	2.81	-.49	4.30**
Mazes	9.19	2.98	8.69	3.14	+.50	3.60**
Verbal IQ	89.61	12.07	89.63	12.13	-.02	.04
Performance IQ	90.16	11.67	89.29	12.22	+.87	.72
Full Scale IQ	88.96	11.35	88.61	11.48	+.35	.13

Note. WISC-R = Wechsler Intelligence Scale for Children—Revised.

^a White M - Black M , uncorrected.

^b $df = 1, 538$.

* $p \leq .10$. ** $p \leq .05$. *** $p \leq .01$.

2.42, $p \leq .01$. As the multivariate F is highly significant, univariate F tests were calculated for each subtest and IQ scale to determine which scores differ significantly and in what direction. The F tests and their significance levels are shown in Table 1. Examination of the separate subtests reveals again that blacks do not earn significantly higher scores on the verbal subtests, either relative to themselves or relative to the matched white sample, contrary to the conclusions of Lesser et al. (1965) and Vance et al. (1979). Whites significantly ($p \leq .05$) exceed blacks on Comprehension, Object Assembly, and Mazes, with a tendency ($p \leq .10$) to exceed also on Picture Arrangement. Blacks significantly ($p \leq .05$) exceed whites on Digit Span and Coding, with a tendency ($p \leq .10$) for higher scores on Arithmetic.

Not only are these results inconsistent with Lesser et al. and Vance et al., they do not support claims of greater cultural bias in the verbal subtests of the WISC-R (Williams, 1974). The WISC-R Information, Vocabulary, and Comprehension subtests are frequently singled out for accusations of blatant cultural bias against blacks. Of these three subtests, only comprehension is relatively more difficult for blacks.

The three subtests on which blacks had their highest levels of performance (Arithmetic, Digit Span, and Coding) form a triad that is frequently referred to in the clinical literature as indicative of "freedom from distractibility." Numerous studies (see Kaufman, 1979, and Lutey, 1977) indicate that performance on these three subtests can be adversely affected by an increase in the subject's anxiety level. Many armchair critics of intelligence testing claim that black children, due to their unfamiliarity with such situations, become inordinately anxious during the administration of an individual intelligence test, and that this anxiety partially accounts for the lower overall scores of these groups. That black children earn their highest scores on those tests that are most sensitive to the effects of anxiety on performance clearly contradicts this claim.

The pattern of differences seen in Table 1 appears somewhat consistent with Jensen's theory of Level I (rote learning and memory) and Level II (complex or transformational cognitive processing) abilities and the gen-

eral finding that white-black differences are greater on Level II than on Level I (Jensen, 1973, 1974; Jensen & Figueroa, 1975). The Arithmetic, Digit Span, and Coding subtests, on which blacks exceed whites, are probably the three most representative measures of level I ability in the WISC-R. Object Assembly, Mazes, Comprehension, and Picture Arrangement are all more closely related to Level II skills.

Another, methodologically initiated, explanation exists to account for the pattern of difference scores, however. Because whites were selected (to match blacks) from a much larger sample, and, to match the black group, whites with relatively low IQs had to be selected, the observed differences could be due to regression to the mean for the whites. One means of correcting this problem would be to have matched subjects using regressed true scores for the white sample. This procedure, however, would destroy the practical aspects of the study that deal with characteristic profiles of blacks and whites with the same *obtained* IQ. Regression effects must nevertheless be investigated. With knowledge of the Full Scale IQ reliability, the obtained means and standard deviations of the Full Scale IQ for the total white sample and our selected subsample, and the correlation of each subtest with the Full Scale IQ, it is possible to calculate regression effects in the present study. To determine what effect the regression problem may have had, regressed means were determined on each subtest for the whites, and the uncorrected difference scores in Table 1 (column *D*) were recalculated with the regressed means. A Pearson correlation was then determined between the reported difference scores and difference scores calculated with the regressed means. The resulting correlation coefficient was .964. This is not surprising, given the WISC-R Full Scale IQ reliability coefficient of .96. Thus, regression effects would have altered the pattern of difference scores only minimally, at most. The magnitude of the effect on individual subtest means was also rather small, with no changes larger than .10 occurring. When significance levels for each of the difference scores are evaluated (using the regressed subtest means) only one change occurs; the Arithmetic subtest moves beyond $p \leq .10$.

Comparisons of Orthogonal Factor Scores

When the intercorrelations among the 12 WISC-R subtests are factor analyzed separately for the white and black samples, the same factors emerge for each group, each with highly similar loadings on the various subtests. Comparison of the factor structure of the WISC-R for the black with the white children from the standardization sample has previously indicated that essentially the same factors, of the same magnitude, emerge for each of the two groups (Gutkin & Reynolds, 1981); other studies comparing WISC-R factor analyses across race for blacks and whites consistently yield similar results (Reynolds, 1982, Note 2). Because of the high congruence of factors in the two samples, indicated by congruence coefficients exceeding .98 for each factor, a factor analysis of the intercorrelations for the combined samples, with $N = 540$, is not only justified, but yields the factor structure of the WISC-R most reliably. Like nearly all other factor analyses of the Wechsler subscales (Matarazzo, 1972, Ch. 11), the present analysis yields three significant group factors, which may be labeled Verbal, Performance (nonverbal), and Memory. The third factor has also been labeled Freedom from Distractibility, but the common cognitive feature of the most highly loaded tests on

this factor (Digit Span and Arithmetic) seems to be short-term memory. Kaufman (1975) recognized this in his comprehensive factor analytic study of the WISC-R, but retained the Freedom from Distractibility nomenclature primarily out of tradition rather than the belief that distractibility is the main source of variance in this factor.

A principal factor analysis (with communalities in the main diagonal) was performed on the intercorrelations of the 12 subscales in the combined samples. The first three principal factors had Eigenvalues greater than 1.00—a common criterion for the significance of factors. The first unrotated principal factor is probably the best estimate of the general or g factor of the battery. The loadings on this g factor are shown in Table 2. Orthogonal rotation of the first three principal factors to approximate simple structure by the well-known varimax criterion, also shown in Table 2, most clearly displays the three group factors of the WISC-R: I, Verbal; II, Performance; III, Memory. Table 2 also shows the communalities (h^2) of each of the variables and the percentage of the total variance accounted for by each factor.

The g factor loadings are relevant to Spearman's (1927, p. 379) hypothesis that the relative magnitudes of mean white-black differences on various tests are directly re-

Table 2
Factor Loadings (decimals omitted) on the General Factor and the Varimax (Orthogonal) Rotated Factors for the Combined White and Black Samples ($N = 540$)

Subtest	Varimax rotated factors					h^2	r^b_{XX}
	g^a	I	II	III			
Information	65	56	15	37	48	85	
Similarities	67	62	22	25	50	81	
Arithmetic	58	27	14	65	53	77	
Vocabulary	78	80	19	26	75	86	
Comprehension	65	60	25	22	47	77	
Digit Span	47	18	14	57	37	78	
Picture Completion	51	33	45	06	32	77	
Picture Arrangement	51	34	36	17	27	73	
Block Design	61	17	67	26	54	85	
Object Assembly	52	18	67	06	49	70	
Coding	30	14	12	29	12	72	
Mazes	43	08	51	19	30	72	
Percentage of variance	32.4	17.7	14.3	10.8	42.9		

^a Unrotated first principal factor.

^b Reliabilities based on standardization sample, from the WISC-R Manual (Wechsler, 1974, p. 28).

lated to the tests' g loadings. Because whites and blacks were intentionally matched on Full Scale IQ in this study, and the Full Scale IQ reflects g more than any other factor, Spearman's hypothesis obviously cannot be completely tested with these data. However, one prediction relevant to these data can be made from the hypothesis, namely, that for white and black samples that are matched on Full Scale IQ there should be a *negative* correlation between the absolute (i.e., unsigned) mean white-black differences on the subtests and their g loadings.¹ A test of this hypothesis is the Pearson correlation between column D (regardless of sign) of Table 1 and column g of Table 2, which is $r(10) = -.67$, one-tailed, $p < .02$. Since both $|D|$ and g are correlated with the subtest reliabilities as given in the WISC-R Manual (last column of Table 2), we should also test this hypothesis after correcting the mean white-black differences and the g loadings for attenuation. This is done by dividing $|D|$ and g for each subtest by the square root of the subtest's reliability coefficient. The Pearson correlation between the corrected $|D|$ and g is $r(10) = -.64$, one-tailed $p < .05$. Thus the one possible prediction from Spearman's hypothesis, for these data, is significantly borne out. One might wonder, however, why the negative correlation, although it is significant, is not larger. There are two possibilities: (a) matching the groups on Full Scale IQ is only a rough approximation to matching the groups on g itself, and (b) whites and blacks also differ on other factors in addition to g . The first possibility is virtually ruled out in these data, by the finding of a correlation of .98 between FS IQ and g factor scores based on the first principal factor in the combined samples. Additionally, the various Wechsler subtests, though differentially related to g , are included on the WISC-R based on the test author's judgment that they are good measures of g and can hardly be considered a random sample of tests with various levels of g saturation. This produces an inestimable restriction of range in the calculation of the correlation between each subtest's g loading and the size of the black-white differences on the subtest. If other tests with smaller g loadings had been included, it is possible

that the correlation could have been significantly larger. More detailed analyses are thus needed to ferret out the actual relationship between g and black-white differences on mental tests. The following analyses are aimed at testing the hypothesis that whites and blacks differ on other factors in addition to g .

Factor scores derived from the three varimax rotated factors, which are uncorrelated except for some minimal dependence due to regression produced in the determination of the factor scores, were calculated for every subject. The factor scores for the combined groups are expressed as standardized z scores, with mean = 0, $SD = 1$. The mean white-black factor scores, SD s, and the mean differences, with tests of significance, on each of the three factors are shown in Table 3. It will probably come as a surprise to many that the F ratios from the analysis of variance show a nonsignificant mean white-black difference on the Verbal factor, but quite significant differences on the Performance and Memory factors. Whites exceed blacks on the Performance factor, whereas blacks exceed whites on the Memory factor. It should be noted that despite their high levels of significance, the racial differences on the Performance and Memory factors amount to less than one-fifth of a standard deviation, equivalent to less than 3 Full Scale IQ points. Such small differences would have some predictive power for the average level of performance of *groups* in tasks that are especially loaded on the Performance and Memory factors but would have no practical interpretive significance for individuals.

It is apparent that there are reliable differences between whites and blacks in cognitive abilities other than g . The true magnitudes of these differences, however, remain to be determined in random (unmatched) national samples of the white and black populations. The nature of the racial difference in the two group factors—Performance and Memory—revealed in this

¹ A negative correlation should result here, because if there are differences due to g , and g has been eliminated by the matching procedure, any remaining differences must be *inversely* related to g , producing a negative correlation.

Table 3

Means, Standard Deviations, and F Tests for Significance of the Difference Between Orthogonal WISC-R Factor Scores of Whites and Blacks

Factor	White		Black		Mean Diff.	F ^a
	M	SD	M	SD		
I. Verbal	+ .018	.870	-.018	.857	+ .036	.23
II. Performance	+ .088	.792	-.088	.851	+ .176	6.22*
III. Memory	-.088	.748	+ .088	.748	-.176	7.45**

Note. Diff. = difference.

^a Degrees of freedom = 1, 538.

* $p < .02$. ** $p < .01$.

study is consistent with certain other findings.

Looking at the rotated factor loadings (Table 2) for Factor II (Performance), we note that the three subtests with the largest loadings are Block Design (.67), Object Assembly (.67), and Mazes (.51). Successful performance on these three subtests probably calls more on spatial visualization ability than any other WISC-R subtests. A number of studies have suggested that blacks perform further below whites on a spatial visualization factor than on any other primary factor, independently of g (Noble, 1978, pp. 327-351; Tyler, 1965, pp. 318-319; for theoretical discussions see Jensen, 1975, 1978; Stevens & Hyde, 1978), though preliminary results from a different line of research tend to provide evidence inconsistent with this hypothesis (Reynolds, McBride, & Gibson, 1981). We consider the white-black differences in spatial visualization ability (independent of g) still an open question, but the present analysis is certainly consistent with the hypothesis of such a difference.

Looking at the rotated factor loadings (Table 2) for Factor III (Memory), we see the largest loadings on Arithmetic (.65) and Digit Span (.57). Successful performance on both of these subtests depends heavily on short-term memory. The simple arithmetic problems, given orally, require the subject to retain all the essential elements of the problem long enough to solve it, and the short-term memory demand of the Digit Span tests is obvious. A large number of studies relevant to Jensen's Level I/Level II theory of abilities and their interaction with race differences consistently shows relatively small or nonexistent differences between

whites and blacks on various tests of short-term memory, even though such tests usually have some moderate g saturation. (For a comprehensive review of this evidence, see Vernon, 1981.) This fact is consistent with the present finding that when memory ability is measured independently of general intelligence, by means of orthogonal factor scores, blacks score higher than whites on the Memory factor. When matched on demographic (but not cognitive) variables, Digit Span is the only subtest failing to display higher mean scores for whites (Reynolds & Gutkin, 1981).

The present results lend support to the Spearman hypothesis that black-white differences are due primarily, but not entirely, to differences in general ability. Spearman's hypothesis cannot account for all ability differences between the races, however. Other ability differences, although of much lesser magnitude than the g difference, occur in the form of a black deficit in spatial-visualization skills coupled with black superiority on tests of rote, short-term memory. Future studies will be designed to estimate the magnitude of these effects in the populations of interest; for now, effects independent of g appear to be small but reliable.

Our results contradict popular views of blacks being disadvantaged on heavily verbal tasks, especially those such as Information and Vocabulary, which are believed to be heavily culture-loaded and specific to the white middle class environment. Of all such verbal subtests that have been criticized, only Comprehension proved more difficult for blacks. Is biased content responsible for this finding? We can hardly accept this explanation, as it would require extrapola-

tion to other tests, leading to the far-fetched conclusion that Arithmetic and Digit Span are somehow biased in content against white children. The anxiety hypothesis of black-white score differences, as in other research (e.g., Reynolds & Gutkin, 1981), also is not supported. The specific sources of black-white differences in mental skills remains to be discovered. It now appears that they are likely to prove even more complex than previously believed.

Reference Notes

1. Jensen, A. R. *The nature of the average difference between whites and blacks on psychometric tests: Spearman's hypothesis*. Unpublished manuscript, 1980. (Available from Arthur R. Jensen, Institute of Human Learning, University of California, Berkeley, California 94720.)
2. Reynolds, C. R. *Test bias: In God we trust, all others must have data*. Paper presented at the meeting of the American Psychological Association, Los Angeles, August, 1981.

References

- Gutkin, T. B., & Reynolds, C. R. Factorial similarity of the WISC-R for blacks and whites from the standardization sample. *Journal of Educational Psychology*, 1981, 73, 227-231.
- Jensen, A. R. Level I and Level II abilities in three ethnic groups. *American Educational Research Journal*, 1973, 10, 263-276.
- Jensen, A. R. Interaction of Level I and Level II abilities with race and socioeconomic status. *Journal of Educational Psychology*, 1974, 66, 99-111.
- Jensen, A. R. A theoretical note on sex linkage and race differences in spatial ability. *Behavior Genetics*, 1975, 5, 151-164.
- Jensen, A. R. Sex linkage and race differences in spatial ability. A reply. *Behavior Genetics*, 1978, 8, 213-217.
- Jensen, A. R. *Bias in mental testing*. New York: The Free Press, 1980.
- Jensen, A. R., & Figueroa, R. A. Forward and backward digit span interaction with race and IQ: Predictions from Jensen's theory. *Journal of Educational Psychology*, 1975, 67, 882-893.
- Kaufman, A. S. Factor analysis of the WISC-R at eleven age levels between 6½ and 16½ years. *Journal of Consulting and Clinical Psychology*, 1975, 43, 135-147.
- Kaufman, A. S. *Intelligent testing with the WISC-R*. New York: Wiley-Interscience, 1979.
- Kaufman, A. S., & Doppelt, J. E. Analysis of WISC-R standardization data in terms of the stratification variables. *Child Development*, 1976, 47, 165-171.
- Lesser, G. S., Fifer, G., & Clark, D. H. Mental abilities of children from different social-class and cultural groups. *Monographs of the Society for Research in Child Development*, 1965, 30(4).
- Lutey, C. *Individual intelligence testing: A manual and sourcebook*. Greeley, Co.: Carol Lutey Publishing, 1977.
- Matarazzo, J. D. *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). Baltimore: Williams & Wilkins, 1972.
- Noble, C. E. Age, race, and sex in the learning and performance of psychomotor skills. In R. T. Osborne, C. E. Noble, & N. Weyl (Eds.), *Human Variation: The biopsychology of age, race, and sex*. New York: Academic Press, 1978.
- Reynolds, C. R. The problem of bias in psychological assessment. In C. R. Reynolds & T. B. Gutkin (eds.), *The handbook of school psychology*, New York: Wiley, 1982.
- Reynolds, C. R., & Gutkin, T. B. Predicting the pre-morbid intellectual status of children using demographic data. *Clinical Neuropsychology*, 1979, 1, 36-38.
- Reynolds, C. R., & Gutkin, T. B. A multivariate comparison of the intellectual performance of blacks and whites matched on four demographic variables. *Personality and Individual Differences*, 1981, 2, 175-180.
- Reynolds, C. R., McBride, R. D., & Gibson, L. J. Black-white IQ discrepancies may be related to differences in hemisphericity. *Contemporary Educational Psychology*, 1981, 6, 180-184.
- Spearman, C. *The abilities of man*. New York: Macmillan, 1927.
- Stevens, M. E., & Hyde, J. S. A comment on Jensen's note on sex linkage and race differences in spatial ability. *Behavior Genetics*, 1978, 8, 207-211.
- Tyler, L. E. *The psychology of human differences* (3rd ed.). New York: Applton-Century-Crofts, 1965.
- Vance, H. B., Hankins, N., & McGee, H. A preliminary study of black and white differences on the revised Wechsler Intelligence Scale for Children. *Journal of Clinical Psychology*, 1979, 35, 815-819.
- Vernon, P. A. Jensen's theory of Level I-Level II abilities: A review. *Educational Psychologist*, 1981, 16, 45-64.
- Wechsler, D. *Wechsler Intelligence Scale For Children—Revised*. New York: Psychological Corporation, 1974.
- Willerman, L. *The psychology of individual and group differences*. San Francisco: W. H. Freeman, 1979.
- Williams, R. L. From dehumanization to black intellectual genocide: A rejoinder. In G. J. Williams & Sol Gordon (Eds.), *Clinical child psychology: Current practices and future perspectives*, New York: Behavioral Publications, 1974.

Received April 5, 1982 ■