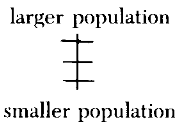
prehension ability is highly loaded with *g* and is widely (but not well) taught in the schools. As an illustration, low-*g* students are initially unable to correctly answer questions such as the following.

In geology the last 11,000 years are called the Recent epoch, and the Recent epoch together with the Pleistocene epoch makes up the Quaternary period. Moreover, the Quaternary together with the Tertiary period makes up the Cenozoic era. The Cenozoic is the only era in which periods are broken down in epochs. The other eras are subdivided only into periods. The era immediately preceding the Cenozoic is the Mesozoic, during which the Jurassic period represents the age of the dinosaurs, although these giant reptiles appeared before the Jurassic and became extinct later than the Jurassic – in the Triassic and Cretaceous periods, respectively. In the still earlier Paleozoic era the first sharks and reptiles appeared during the next-to-last period, the Carboniferous, while in the last period of this era, the Permian, reptiles flourished. Preceding the Carboniferous period was the Devonian, and before that, from earliest to latest, the Cambrian, Ordovician, and Silurian periods. Write the 11 periods in order from earliest to latest on a diagram. Do not write eras or epochs.

However, their performance improves greatly after using TAPS while working through a series of 60 problems, beginning with easy ones like this (Whimbey 1983).

Atlanta has a larger population than Birmingham but a smaller population than Chicago. Write the names of the three cities in order on the diagram.



Significant gains have been made on a standard reading measure, the Iowa Silent Reading Test, but an evaluation of the practical, long-term impact will take several more years (Whimbey 1981).

In closing I would like to draw attention to a few additional research questions raised by Jensen's findings. If blacks have slower reaction times, how have they come to dominate boxing and excel in other sports like baseball? If they can't get their finger off the button of the reaction-time apparatus as quickly as their white counterparts, how do they duck their punches and hit their pitches so well? The answer is not simply superior muscular strength, because both American and international weightlifting is dominated by whites. Nor is it muscular coordination, since Jensen's Figure 6 shows no difference between blacks and whites here. Aside from athletics, the reaction-time research seems at odds with the prominence of blacks in the creation and performance of jazz, some of which (for example, that of Thelonius Monk) is rich, complex, and sophisticated. As Jensen suggests, much research is still needed on *g* and other human abilities.

Jensen's support for Spearman's hypothesis is support for a circular argument

James R. Wilson
Institute for Behavioral Genetics, University of Colorado, Boulder, Colo. 80309

Having received an earlier draft of the target article, we prepared and submitted a paper commenting on several aspects of it, and we included new analyses from the Hawaii Family Study of Cognition that were relevant to the argument (Nagoshi et al., in press). We furnished Professor Jensen with a copy of this manuscript; however we unfortunately see no indication in his article that he has considered the arguments or data presented therein.

I would like to reiterate here one of the arguments we presented: "Because a group difference on *g* requires group differences on tests which load on *g*, an observed group difference in general mental ability may necessarily result in a correlation between group differences on individual tests and their *g* loadings" (Nagoshi et al., in press). Another way of saying this is that, given a substantial group difference on *g* (such as is commonly reported for blacks vs. whites), it is hardly surprising that there will be a substantial group difference on those tests which load most heavily on *g*, since they in a very real sense *define g*. Whether it is *g* that we conceive to be theoretically prior or the actual tests hardly matters; we have but one phenomenon (the group difference), and we add nothing to our understanding of the phenomenon by running the argument around in a circle.

Author's Response

The black–white difference in *g*: A phenomenon in search of a theory

Arthur R. Jensen
School of Education, University of California, Berkeley, Calif. 94720

The 29 commentaries present such a diversity of opinions and observations on so many different aspects of the target article as to make it virtually impossible to do justice in my response to every single point. It will be necessary to focus on those issues that show some communality among commentators or that raise questions that seem most central to the main findings and are most apt to help clear the way for further research and theoretical formulations. I will try, however, to touch upon as many of the points raised by the commentators as possible, even if it means adopting a fairly telegraphic style, with abrupt changes of topic.

Most of the comments fall into one of two main categories: (1) Spearman's hypothesis per se and the psychometric and statistical problems surrounding it, and (2) the relation of response latency, or reaction time (RT), on a variety of elementary cognitive tasks (ECTs) to psychometric *g* and associated methodological and interpretive issues.

Spearman's hypothesis per se

There is rather less explicit agreement or disagreement with Spearman's hypothesis than I had expected, given that it was the central theme of my target article. The test of Spearman's hypothesis is the significant and consistent (across 12 studies) correlation between psychometric tests' *g* loadings and the magnitudes of the mean black–white differences on the tests (expressed in standard score or σ units). Nine of the commentators (**Brand, Cattell, Eysenck, Gordon, Jones, Kline, Nettelbeck, Nichols,** and **Stanovich**) explicitly regard the hypothesis as having been borne out by the evidence. Two (**Gustafsson** and **Baron**) express doubts or propose a counterhypothesis. Three (**Johnson & Nagoshi, Schönemann,** and **Wilson**) seem to accept the hypothesis as borne out, but claim that this outcome was inescapable, being a mathematical artifact or a "circular argument" preordained by the workings of factor analysis. The remaining 13 commentators express no opinion one way or the other regarding Spearman's hypothesis per se.

Sternberg, however, does not quite fall into any of these categories. He claims that the analysis merely restates what we already knew, namely, that blacks score lower than whites on conventional intelligence tests. That fact has indeed been well known for a long time. But that is not the issue specifically addressed by Spearman's hypothesis, which arose in the first place from the observation that there is considerably more to the black–white difference on psychometric tests than just the overall difference itself, the important point being that the magnitude of the black–white difference *varies* across different tests. I have not found any thorough examination of this phenomenon anywhere in the previous literature. There would be little if any scientific leverage in observing still one more black–white difference on one more test. What has not been discussed, much less understood, is the *variation* of differences, which, if it proves to be a reliable phenomenon, could provide some leverage for further understanding the nature of the black–white differences in cognitive performance. Testing Spearman's hypothesis, or investigating the variation among differences, is an essential step toward an adequate account of the black–white differences on psychometric tests. Sternberg's belittling of this aim is surprisingly unanalytical for an otherwise generally very analytical psychologist. Attempting to reduce these findings to nothing more than the well known average difference of about 1 σ on "conventional intelligence tests" not only misses the essential question that gave rise to Spearman's hypothesis, but also tars such research with the popular opprobrium attached to IQ tests. Is it not a reasonable question to ask (assuming we are interested in the subject at all) which content features or psychometric characteristics of tests are associated with the conspicuous variation in the size of the mean black–white difference on different tests? Might not such inquiry afford clues as to the essential nature of the black–white difference, or at least point investigators in the best direction for further study? What the present analysis consistently shows is that variation in the black–white difference is not systematically associated with such surface or content characteristics of tests as whether they are verbal or nonverbal, culture-loaded or culture-reduced, performance or paper-and-pencil, pictorial or figural, and so on but is most consistently associated with a latent trait, *g*, or the largest common factor in virtually any sizable battery of diverse cognitive tasks. The nature of the black–white difference, therefore, must be sought in the nature of *g* rather than in the intellectual content and other surface features of conventional psychometric tests.

The inevitability–circularity–artificiality claim. Several commentators regard the outcome of testing Spearman's hypothesis as inevitable or artifactual or a circular argument. **Jones** believes that any other conclusion from the results would be totally unexpected. If it is unexpected to Jones, it is largely because Jones, a sophisticated psychometrician who has investigated black–white differences, already knows the kinds of tests that show the largest differences and the fact that constructors of conventional intelligence tests select item types that are *g*-loaded. They select such items not necessarily because they are *g*-loaded but because it is found that the most *g*-loaded items maximize predictive validity for the kinds of practical criteria for which tests are commonly used. The literature on group differences in test scores attributes differences almost exclusively to specific contents and surface features of tests, and the demonstration of what Jones regards as totally expected (i.e., the substantiation of Spearman's hypothesis) actually contradicts the conventional and popular view of black–white test differences. Moreover, not all group differences on a battery of psychometric tests are *g* differences, as I showed in the comparison of preverbally deaf children and normal-hearing children. The correlation between WISC-R subtest group differences and subtest *g* loadings for deaf and hearing children was in fact negative–the opposite of the black–white comparison. Jones assumes that this "unexpected" finding must be due to a different pattern of *g* loadings for deaf and hearing children. Yet Braden (1984, p. 406) has reported a congruence coefficient of +0.988 between the *g* factor loadings of the deaf and hearing groups, that is, virtual identity of the *g* factor across these groups. But the profile of group differences on the subtests is negatively correlated with the profile of the subtests' *g* loadings. True, the overall hearing–deaf difference is only about one-fifth as large as the typical black–white difference. But that cannot be the cause of this outcome. I have shown that the effect of inbreeding depression is to lower the WISC IQ just about as much as Braden reported for the effect of deafness on the Performance IQ. Yet the varying effects of inbreeding depression on the WISC subtests are correlated about +0.80 with the subtests' *g* loadings (Jensen 1983a).

As for **Wilson**'s claim of circularity, it is his own argument, not Spearman's hypothesis, that is circular. Of course, if one postulates (as does Wilson) that a group difference is mainly a *g* difference, then it is indeed inevitable that the group differences on various tests will be correlated with the tests' *g* loadings. One can always make a proposition circular by stating the conclusion in the premises. The same fallacy is voiced by **Johnson & Nagoshi,** who, in their first sentence, state that "any group difference in *g* would of necessity be reflected in the tests that load on *g*." This is of course a mere tautology. Change the statement to "any group difference in IQ (or total score, etc.)" and it is no longer a tautology or inevitability. After stating the tautology, Johnson & Nagoshi claim that "his finding in itself casts serious doubts on the validity of Jensen's conclusions concerning black–white differences in cognitive abilities." But this claim is a non sequitur. Do Johnson & Nagoshi mean to imply that this tautology contradicts Spearman's hypothesis? After their puzzling first paragraph, Johnson & Nagoshi go on to show some other theoretically interesting relations between *g* and certain familial and social variables in their own study of various populations in Hawaii, and one could hardly disagree with their concluding statement that "there is clearly a need for even more basic research on the nature of *g*."

Schönemann illustrates the same kind of tautology mathematically, showing that if one "builds in" a large enough difference between groups on a number of correlated variables (which thereby yield a general factor), the groups will differ on the general factor. It appears to me that this is another case of stating the premises or conditions necessary for a given outcome. I do not see that it differs essentially from saying, for example, that if two cars start a race from the same point and traverse the same distance, the car with the faster average speed will cross the finish line ahead of the car with the slower average speed. But do the premises that the average speeds are different and that the distance is the same make the observation that one car arrives at the finish line ahead of the other merely an artifact or an illusion? On the other hand, one can point to many conjectures by psychologists in the literature on black–white IQ differences that are contradicted by the very conditions or premises that Schönemann demonstrates as sufficient for Spearman's hypothesis, for example, equal covariance matrices and large enough differences on the mean vectors. But Schönemann's demonstration apparently leads him to agree with a false, or at best theoretically too limited, conclusion, namely, the statement he quotes from my *Bias in Mental Testing* (1980a). Although the conditions stated therein *could* produce the appearance of Spearman's hypothesis, these conditions are neither necessary nor sufficient to account for the actual findings. Since 1980, I have explicitly investigated this matter, and I find that neither the variation in the *g* factor nor the varying magnitude of the black–white difference on various tests is at all dependent on differences in test reliability or on variation in item or subtest difficulty level. High and low *g*-loaded tests, even when perfectly matched on reliability, still show large and small black–white differences, respectively. Moreover, we have found that different single items of the Raven Progressive Matrices test can differ in their *g* loadings even when they are perfectly matched on item variance [i.e., $p(1 - p)$, where p is proportion passing]; the more complex (hence more difficult) items are generally the more *g*-loaded, even when $p(1 - p)$ is the same for the simple and complex items (e.g., $p = .80$ and $p = .20$). In brief, *g* can vary independently of reliability and range restriction, even among tests or items that are quite homogeneous in form and content. It has become increasingly clear in recent years that neither *g* nor the black–white difference on cognitive tests is merely a psychometric artifact.

Baron is right in noting that the reliability of a test can affect both its *g* loading and its power to discriminate groups. But this does not mean that Spearman's hypothesis depends on differences in test reliability, although such differences could conceivably simulate an outcome consistent with the hypothesis when the hypothesis was actually false. However, the present results cannot be explained in this way, as I have already shown in the target article. When the *g* loadings and black–white differences ($\bar{D}$) are corrected for attenuation, Spearman's hypothesis still holds (see Table 3 in the target article). The lowering (by about .10) of the correlations between *g* and $\bar{D}$ is adequately explained by the greater restriction of range of the disattenuated *g* loadings. The use of parallel-forms test–retest reliabilities rather than internal-consistency (split-half or K-R [Kuder-Richardson] 20) reliabilities would be a nice addition but would be most unlikely to alter the results appreciably. Although the two forms of reliability are clearly distinct conceptually, empirically the resulting reliability coefficients (r_{xx}) generally run quite parallel. I think that Baron makes too much of the partial correlation, in which r_{xx} is partialed out of the zero-order correlation between *g* and $\bar{D}$. Although the resulting partial correlation is capable of testing the null hypothesis, beyond that, its actual magnitude, unlike the correction for attenuation, cannot be interpreted as yielding a closer approximation to the true correlation between *g* and $\bar{D}$. There is, of course, no demonstration of an inherent theoretical connection between a test's parallel-form retest reliability and either its true (i.e., disattenuated) *g* loading or its true discriminability between populations. Reliability and *g* are certainly not the same construct, even though in some test batteries they may be adventitiously correlated. Reliability is largely a function of test length. The Digit Span subtest of the Wechsler scales, for example, has one of the lowest reliabilities in the Wechsler battery and also one of the lowest *g* loadings, and it shows one of the smallest black–white differences of any subtest. In a number of my previous studies, however, I have used repeated parallel forms of the forward Digit Span test to increase the reliability of the composite Digit Span test up to values above .90, that is, as high as the reliability of the Full Scale IQ. Even so, the *g* loading of Digit Span is still much lower than the *g* loadings of, say, Vocabulary and Block Design. Also, the size of the black–white difference on the highly reliable forward Digit Span test is still among the smallest of the differences on any of the many tests we have used in our research (e.g., Jensen 1971; 1973b; 1974a; Jensen & Figueroa 1975; Jensen & Innouye 1980).

Gordon has made a striking contribution to the methodology of testing Spearman's hypothesis, based on the equivalence of the congruence coefficient (or index of factor similarity) and the correlation between factor scores. The point biserial correlation (r_{pb}) of test scores with the black–white dichotomy is clearly equivalent to the tests' loadings on a black–white factor. The question then is whether factor scores based on this black–white factor were computed for every subject, and if factor scores based on the *g* factor (the first principal component) of all the tests in a given battery were computed for every subject, the correlation between the two sets of factor scores—the black–white factor scores and the *g* factor scores—would be equal to the coefficient of congruence between the tests' loadings on the black–white factor and the loadings on the *g* factor. (Although this equivalence would hold exactly only for a *g* factor computed as the first principal component, and the present analyses are based on the first principal factor or on the Schmid-Leiman second-order *g*, these are only negligibly different from the first principal component in the present data sets. Therefore, Gordon's figures would probably differ only in the third decimal place.) The congruence coefficients shown in Gordon's Table 1 range between .915 and .993, with an average of about 0.97, that is, an almost perfect correlation between factor scores based on *g* and the magnitude of the black–white difference, as Gordon concludes. This is a striking substantiation of Spearman's hypothesis, albeit an inferential substantiation, based on the correctness of Gorsuch's (1974, p. 253) claim of equivalence between the principal component factor score correlation and the congruence coefficient. For those who would like to see a precise empirical demonstration of this outcome as well, the total Wechsler (WISC-R) standardization data from the study by Jensen and Reynolds (1982) are available and can be subjected to a direct determination of the equivalence of the factor score correlation and the congruence coefficient. This analysis will be done as soon as feasible, and a note on the results will be submitted to a forthcoming Continuing Commentary section in this journal.

Factor analysis and the nature of *g*. In the target article I tried to treat *g* as empirically as possible, without bringing in any particular theory of *g* or allowing subjective judgments or theoretical preconceptions to determine the *g* factor or its relation to the black–white difference. As a starting point, I thought it best to take whatever *g* the available data sets yielded by an objective method of analysis, even though some of the available test batteries were rather far from representing an ideal sampling of the whole domain of abilities measured by psychometric tests. Never was a test battery included or excluded because of how well the particular collection of tests conformed to any particular theoretical conception of the "ideal *g*," whatever that might mean. I agree with the observation of **Gustafsson, Jones,** and **Vernon** that the *g* factors extracted from these 11 quite diverse test batteries are bound to vary to some degree, which cannot be precisely determined from these data. As correctly noted by **Kline,** however, the fact that the *g* factor varies somewhat according to the different compositions of these batteries could only attenuate the test of Spearman's hypothesis. Yet the hypothesis was borne out by every battery. **Gustafsson** notes that generally in these particular batteries the tests with the largest *g* loadings and largest black–white differences are of the achievement-laden type frequently characterized as crystalized *g*, or g_c, as contrasted with fluid *g*, or g_f. But it might well be that in culturally or educationally homogeneous populations (as indicated, for example, by their high similarity in factor structure), verbal and achievement-type tests yield even better measures of g_f than the often less reliable and spatially loaded tests most commonly used to represent g_f. The *g* of most of the test batteries used in this study is undoubtedly some amalgam of g_c and g_f. But if these batteries could be subjected to a hierarchical or Schmid-Leiman factor analysis along with a much larger collection of tests that sampled more widely the entire psychometric domain, I think it would be a safe prediction that the topmost *g* of the hierarchy (call it Spearman's *g*) would be larger (in variance accounted for) than g_c or g_f or the two combined and that the residualized g_f would be reduced to practically nil, most of it being absorbed by Spearman's *g*. Recent hierarchical factor analyses of test batteries with broad samplings of abilities have shown exactly this picture (Gustafsson 1984; Undheim 1981a; 1981b; 1981c). Spearman's *g* and g_f are either very similar or the same, and much of the variance of the kinds of tests that are usually most heavily loaded on g_c is absorbed into the top hierarchical *g* when residualized by the Schmid-Leiman procedure. Hence one cannot accept as a cogent criticism Gustaffson's comment that my analysis leaves Spearman's hypothesis largely uninvestigated. However, it would be very desirable to see Spearman's hypothesis tested using the broad sample of tests that, in Gustafsson's (1984) own study, yielded what he might consider an "ideal" *g* and led him to conclude that *g* is identical with g_f. This conclusion led Gustafsson (1984) to a most important observation: "Formulated in simple terms this result implies that scores obtained on a test consisting of the broadest and most representative sample of tasks are virtually perfectly correlated with scores obtained on a small set of g_f tasks. The most interesting question must then be why the g_f tests have such power of indexing general intelligence" (p. 195).

This, I think, is the most telling criticism of Humphreys's purely descriptive definition of general intelligence, a conception that **Jones** seems to advocate that I should adopt. Humphreys (1971) has defined general intelligence as follows:

> Intelligence is defined as the entire repertoire of acquired skills, knowledge, learning sets, and generalization tendencies considered intellectual in nature that are available at any one period of time. An intelligence test contains items that sample the totality of such acquisitions. The definition of intelligence here proposed would be circular as a function of the use of intellectual if it were not for the fact that there is a consensus among psychologists as to the kinds of behaviors that are labeled intellectual. Thus, the Stanford-Binet and the Wechsler tests can be considered examples of this consensus and define the consensus. (Pp. 31–32)

My own reservations about this definition have been expressed in detail elsewhere (Jensen 1984c). The definition is essentially theoretically barren. In relation to the earlier quotation by Gustafsson, it is a theoretically crucial fact that intelligence, as defined by Humphreys, can actually be measured adequately by a limited number of tests that involve much less than the totality of the repertoire of acquired skills described by Humphreys. One does not need to sample from the totality of this repertoire in order to measure its general factor. In fact, it is now beginning to appear that one may need to measure only certain aspects of the averaged electrical potentials of the brain elicited by auditory "clicks" (Hendrickson & Hendrickson 1980). Humphreys's definition deals only with what **Eysenck,** following Hebb, has termed Intelligence B, which comprises the multifarious manifestations of Intelligence A, characterized by Eysenck as a "capacity of the central nervous system and cortex to process information correctly and without error." There is nothing in the Humphreys definition that would lead one to expect the existence of a *g* factor in the varied repertoire described by his definition or to imagine that the same *g* factor could be measured by tests tapping very different contents of the repertoire–the important phenomenon referred to by Spearman (1927) as "the indifference of the indicator" of *g*.

As an example of this phenomenon, I cited the fact that the *g* factors extracted separately from the Wechsler verbal subtests and the performance subtests are correlated .80 with each other, despite the highly dissimilar contents of the verbal and performance tests. **Vernon** appears to cast doubt on this claim by citing a correlation of .67 between the Verbal and Performance IQs of the Wechsler Intelligence Scale for Children–Revised (WISC-R). I haven't determined the correlation between the *g* factors of the Verbal and Performance subtests of the WISC-R; my statement was based on this determination for the Wechsler Adult Intelligence Scale (WAIS), in which even the simple correlations between the Verbal and Performance IQs range between .77 and .81 for various age groups (Matarazzo 1972, p. 243). Clearly, very dissimilar test batteries yield very similar *g*s. Is the *g* of all the Wechsler subscales mainly g_c, as **Cattell**'s statement suggests, or does it also represent g_f to a substantial degree? One might predict from Gustafsson's (1984) observations that a *g* extracted from such a diverse battery as the Wechsler would most probably come close to Cattell's g_f. Raven's Matrices, like Cattell's Culture Fair Tests of g_f, is generally considered a quintessential test of g_f. It is therefore noteworthy that when the Raven Matrices (Advanced) was factor-analyzed among the 12 WAIS subtests, it showed a higher *g* loading (+0.80) than any of the WAIS subtests; Block Design, Vocabulary, and Arithmetic were next in order, with *g* loadings of +0.69, +0.64, and +0.64, respectively (P. A. Vernon 1983).

The robustness of *g* across diverse test batteries was shown long ago in a study by Garrett, Bryan & Perl (1935) who factor-analyzed a battery of six varied memory tests (meaningful prose, paired-associates, free recall of words, digit span, memory for forms, and memory for objects) and extracted the general factor. This battery of tests was then factor-analyzed along with four other diverse tests not especially involving memory (motor speed, vocabulary, arithmetic, and form board). The *g* loadings of the memory tests in the two analyses correlated .80. The overall correlation between *g* factor scores based on just the memory tests and *g* factor scores based on just the nonmemory tests was .87. This is evidence that the *g* of the six memory tests is very close to the *g* of the nonmemory tests. To be sure, the memory tests were not as highly loaded on *g* (average *g* loading = .42) as the vocabulary and arithmetic tests (average *g* loading = .65), but what little *g* the memory tests have is much the same *g* as found in the nonmemory tests. One would like to see larger-scale studies of this type based on many diverse psychometric tests to determine the variance of correlations between *g* factor scores extracted from different nonoverlapping sets of tests, controlling for reliability.

A set of data provided by R. T. Osborne (personal communication) but not used in the target article, since it is unpublished data, lends support to **Cattel**'s conjecture that, when g_c and g_f can be clearly distinguished by including in the factor analysis a large enough number of the types of tests that will permit the emergence of these two factors, the tests' loadings on g_f would be more highly correlated with the black–white differences than the loadings on g_c. Osborne's battery included seven of the most "fluid" tests from the Educational Testing Service's "Kit of Reference Tests for Cognitive Factors" (French, Ekstrom & Price 1963) (Cube Comparisons, Identical Pictures, Formboard, Surface Development, Spatial, Paper Folding, and Object Aperture). The "crystalized" tests in the battery were the Calendar Test, Arithmetic, the Wide Range and Heim Vocabulary Tests, and Spelling. All 12 tests were given to 608 white and 246 black urban school children. Factor analyses with varimax rotation, performed separately in each group, yielded two orthogonal (i.e., uncorrelated) factors clearly identifiable as g_f and g_c, both of which showed high congruence between the black and the white samples. The Spearman hypothesis was examined separately for g_c and g_f. The correlation between tests' g_c loadings and the mean black–white differences is −0.24 for white g_c loadings and −0.02 for the black; neither *r* is significant. The correlation between loadings on g_f and the black–white difference is +0.56 ($p < .05$) for whites and +0.42 ($p < .10$) for blacks. Thus, the mean black–white differences on these 12 tests are more highly related to the tests' loadings on g_f than on g_c. This result seems to contradict the popular belief that the black–white difference on tests largely involves differences in scholastic learning as characterized by the "crystalized" component of variance in test scores. There is some ambiguity in this study, however, owing to the fact that virtually all the nominal g_f tests are also known to involve spatial visualization ability (g_v) as well as g_f; since nonspatial fluid tests were not included, g_f and g_v could not be distinguished, and so what appears as g_f is actually some amalgam of g_f and g_v. How closely the black–white difference is associated with each of these components separately is not known.

Another study (Jensen 1973b) of large representative samples totaling about 200 white, black, and Mexican-American Californian school children used 17 tests which included nonspatial as well as spatial tests of "fluid" ability (Lorge-Thorndike Nonverbal IQ, Raven Matrices, Figure Copying), three short-term memory tests, and typical "crystalized" tests (Lorge-Thorndike Verbal IQ and the Stanford Achievement battery of seven scholastic achievement tests). A number of socioeconomic indices (Gough Home Index) were also included. Varimax factor analysis yielded four orthogonal factors corresponding to g_c (Verbal IQ and Achievement Tests) and g_f (nonverbal tests), as well as a rote memory factor and a socioeconomic status factor. The mean factor scores of each of the populations on each of the factors are shown in Figure 1. The black–white difference in mean factor scores scarcely differs between the g_c factor (verbal IQ and achievement) and the g_f factor (nonverbal IQ). It should be noted that these are uncorrelated factors. This and other evidence, I believe, drastically undermines **Gustafsson**'s criticism that the differing compositions with respect to g_c and g_f of the various test batteries used to test Spearman's hypothesis has resulted in the hypothesis's remaining largely untested.

Jones cites an article (Jones 1984), which I have not yet seen, showing that "the average scores of the nation's black students on aptitude and achievement tests have steadily risen, relative to average scores for white students, over the past 15 years." The basis for this claim will have to be reconciled somehow with the recently announced results of the Armed Services Vocational Aptitude Battery (ASVAB), a set of ten aptitude and achievement tests administered to a large national probability sample representative of American youths ages 16 to 23 years (Office of the Assistant Secretary of Defense 1982). The mean black–white differences (in standard score units) on some of the ASVAB scholastic achievement tests are Arithmetic Reasoning 1.16, Word Knowledge 1.30, Paragraph Comprehension 1.08, and General Science 1.23. These differences are at least as large as the black–white difference on the Army Alpha at the time of World War I or on the Army General Classification Test in World War II. If there is a genuine discrepancy between

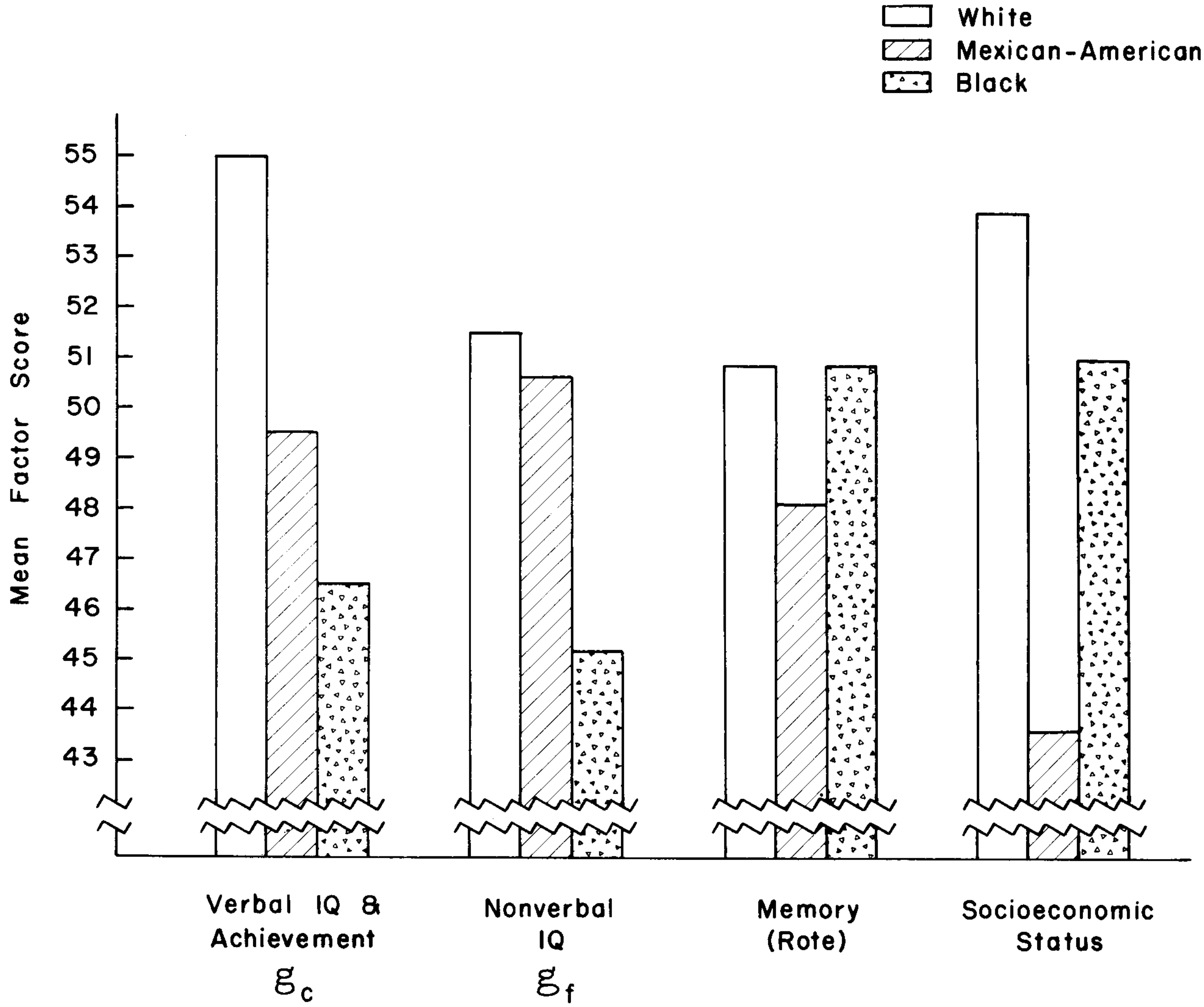


Figure 1. (Response). Mean factor scores (mean = 50, σ = 10 within each grade level) for four variables, comparing white, black, and Mexican-American samples in grades 4, 5, and 6. The factor scores are orthogonal; that is, the scores on any one factor reveal differences between subjects who are statistically equated on the three other factors. (From Jensen 1971, Table 6.)

Jones's test results and the recent ASVAB test results, the discrepancy may be at least partly explainable in terms of Spearman's hypothesis; that is, the ASVAB tests may be more highly *g*-loaded than Jones's tests.

Borkowski & Maxwell claim that although a relationship between tests' *g* loadings and the size of black–white differences has been demonstrated, it has not been shown that the black–white difference is predominantly a difference in *g*, and hence the "weak" form of Spearman's hypothesis remains untested. They have apparently overlooked the study by Jensen and Reynolds (1982) that explicitly apportions the total between-group (black–white) variance to each of the orthogonalized hierarchical factors that emerged from a Schmid-Leiman factor analysis of the WISC-R. This study, based on the national standardization sample of the WISC-R (1868 whites and 305 blacks), showed that the black and white groups differed significantly in mean factor scores on all four of the common factors extracted from the WISC-R: *g*, verbal, performance, and memory. But in terms of the total variance between groups accounted for, the *g* factor accounted for more than seven times as much intergroup variance as the other three factors combined. The four common factors together contribute 89% of the total intergroup variance; the remaining 11% is due to the specificity of the 13 subtests. The same kind of analysis, which was based on factor scores for every subject, was impossible in the ten other studies, for which the scores of individuals were not available. The weak form of Spearman's hypothesis, however, could be further investigated in these studies by including in the test intercorrelation matrix the point-biserial correlations of the black–white dichotomy with each of the tests and then factor-analyzing the matrix to see precisely the magnitudes of the loadings of the black–white variable on each of the orthogonal factors extracted from the matrix. When this analysis is done with the WISC-R data, the results, of course, are completely consistent with those I have just reported, showing the black–white variable to have by far the largest loading on *g*. It is hard to imagine that very different outcomes would be found in the ten other test batteries, but in order to leave no doubts about the answer to this question, I will do the required factor analyses and report the results in Continuing Commentary.

Animal intelligence. It is difficult to evaluate **Macphail**'s claim that there is nothing resembling *g*, or individual differences in intelligence, either between or within different species of nonhuman vertebrates. Any behavioral differences that might be interpreted as differences in cognitive ability or in some general capacity for dealing with complexity, it seems, can also be attributed to species differences in specific sensory and motor capacities or to differing instincts and drives. The literature on comparative psychology, I believe, leaves much room for doubting Macphail's claim, although the null hypothesis, which Macphail seems to favor, may be difficult to reject definitively at present. The main problem is one of devising tests that are deemed equally appropriate across species which differ widely in sensory and motor equipment and in appetites and instinctual behaviors. The problem will have to be debated and resolved empirically, if possible, by experimental comparative psychologists and ethologists. The speed of acquisition of learning sets has been found to be related to intelligence in humans (Hunt 1961, p. 83) and also shows clear inter- and intraspecies differences. As Harlow (1959) has observed, "All existent LS [learning set] data on all measured species are in keeping with the anatomical data bearing on cortical complexity, and it is obvious that LS techniques are powerful measures for the intellectual ordering of primate and possibly even nonprimate forms" (p. 507). Interspecies differences in complexity of behavioral capacities are related to brain size (in relation to body size) and to the proportion of the brain not involved in vegetative or autonomic and sensorimotor functions. According to Jerison (1973), development of the cerebral cortex, the association areas, and the frontal lobes parallels species differences in behavioral complexity. It has been found that the tests which have shown differences in problem-solving capability between monkeys and apes, and even individual differences between chimpanzees, have shown the same rank order of difficulty when they are given to human children as when they are given to apes; this suggests that the tests involve similar capacities across species (Viaud 1960, pp. 44–45).

Macphail harks back to Spearman's (1923, p. 346) original notion of *g* as a kind of "mental energy." Although Spearman intended this description merely as an analogy or metaphor, the notion still has intuitive appeal. High-*g* persons actually give the appearance of possessing more spontaneous mental energy, which they bring to bear on almost everything they do of a cognitive nature, and they also seem to be more persistently active in cognitive ways. But these characteristics may only be the by-products of their greater speed and efficiency of information processing. Equating *g* with drive, formulated as Hull's "big *D*," as suggested by Macphail, would seem to run into difficulty with the Yerkes-Dodson law (Yerkes & Dodson 1908), which is the now well-established empirical generalization that the optimal level of drive (*D*) for learning or performance of a task is inversely related to the degree of complexity of the task; that is, a *lower* level of *D* is more advantageous for the performance of more complex tasks. In this respect, *D* is just the opposite of *g*. The *g* loading of tasks increases with task complexity, and persons who score highest in the most *g*-loaded tests are more successful in dealing with complexity. From what research has taught us about Hull's *D* and the Yerkes-Dodson law, one would not predict high-*D* persons to perform like high-*g* persons as a function of task complexity. In humans, changes in drive and arousal are reflected in pupillary dilation. Ahern and Beatty (1979) measured the degree of pupillary dilation as an indicator of effort and autonomic arousal when subjects are presented with test problems. They found that (1) pupillary dilation is directly related to level of problem difficulty (as indexed both by the objective complexity of the problem and the percentage of subjects giving the correct answer) and (2) subjects with higher psychometrically measured intelligence show less pupillary dilation to problems at any given level of difficulty. (All subjects were university students.) Ahern and Beatty concluded:

> These results help to clarify the biological basis of psychometrically-defined intelligence. They suggest that more intelligent individuals do not solve a tractable cognitive problem by bringing increased activation, "mental energy" or "mental effort" to bear. On the contrary, these individuals show less task-induced activation in solving a problem of a given level of difficulty. This suggests that individuals differing in intelligence must also differ in the efficiency of those brain processes which mediate the particular cognitive task. (P. 1292)

Unitarianism versus componentialism. Questions are raised by both **Brand** and **Nichols** concerning whether *g* variation has unitary or multiple causation, and to what extent it arises from polygenic effects or from correlated environmental influences. These questions are also implicit in several other commentaries. They are really the crux of current theorizing about *g*. These issues are simply unresolved at present, but progress is being made. I do not see a sufficient empirical basis as yet for predicting whether the physiological substrate of *g* will eventually turn out to be some "unitary" feature of neural activity (e.g., cortical conductivity, speed of synaptic transmission, number of neurons, amount of branching, number or organization or complexity of cell assemblies, or capillary blood supply to the cortex) or the resultant of many such features. The well-established fact of the genetic heritability of *g*, however, makes it virtually certain that some substantial proportion of the *g* variance must ultimately find explanation at the neurophysiological level. Cognitive componential theory in all its contemporary forms represents a different level of analysis; it is a behavioral analysis of various cognitive tasks in terms of a limited number of abstracted information processes, or "components," having the status of intervening variables or psychological constructs that are hypothesized to mediate or execute different cognitive tasks. These hypothesized components, or information processes, are operationally definable, and individual differences in them are measurable, at least indirectly, by means of various chronometric techniques. The *g* yielded by factor analysis of psychometric tests, according to the componential view, results from there being certain elementary cognitive processes (and perhaps also metaprocesses) that are required for successful performance on virtually all test items. But measures of the elementary cognitive tasks are themselves intercorrelated, and when factor analyzed they yield a *g* that is correlated with the *g* of psychometric tests. Hence there is a kind of infinite regress of task intercorrelations getting at essentially one and the same *g*, at times more or less obscured or attentuated by task specificity and measurement error. At the very end of this regress of *g* across levels of analysis, presumably, is some physiological substrate, the nature of which is still highly speculative. But we will probably not have a scientifically satisfying explanation of *g* until *g* has been clearly linked to its biological structures or physiological mechanisms. This field is wide open for theoretical speculation and empirical investigation. I do not rule out the possibility, favored by **Brand**, that the basis of *g* at this level could be something much simpler than what we can observe at the psychological or behavioral level of analysis, just as the basic cause of a disease is often much simpler than its multifarious symptoms.

Evoked potentials and *g*. One cannot deny **Callaway**'s assertion that brain electrical potentials, or evoked potentials, are not *necessarily* correlated with intelligence. **Carlson** expresses similar caution at this stage of this research. It is one of the primary aims of current research in this field to discover the specific procedural conditions that will yield the most substantial correlations between certain aspects of the average evoked potential (EP) and psychometric *g*. A recent study by Haier, Robinson, Braden & Williams (1983), for example, has identified various experimental conditions and methods of measurement that have resulted in some of the inconsistent findings in this field. Haier et al. identify those particular conditions that show the highest correlations between EP and IQ. They conclude:

> Perhaps, the most startling conclusion suggested by this body of work is not just that there is a relationship between brain potentials and intelligence, but that the relationship is quite strong. This supports the proposition that the variance of intelligence, with all its complex manifestations, may result primarily from relatively simple differences in fundamental properties of central brain processes. (P. 598)

Schafer's comment provides further striking evidence of the relation between certain parameters of the EP and psychometric *g*. Not only do his data show an overall multiple correlation of +0.64 (or +0.80 corrected for restriction of IQ range in his sample) between the EP parameters and the WAIS Full Scale IQ, but more importantly they also show that the degree to which each of the 11 subtests loads on the *g* factor is directly related to the degree of each subtest's correlation with the EP. Figure 2 shows this relation for the EP habituation index, as defined by Schafer. (The *g* factor here is estimated by the first principal component, provided by Schafer.) Correcting the correlation for attenuation with the reliabilities of the WAIS subtests in the standardization sample results in a lowering of the correlation in Figure 2 from +0.897 to +0.891. Partialing out the subtest reliabilities produces exactly the same result for these data. Moreover, this is not an isolated finding. Eysenck and Barrett (1985), measuring a different parameter of the EP, reported a correlation (Spearman's rho) of +0.95 between WAIS subtests' *g* loadings and the subtests' correlations with the EP measure. It is probably more than sheer coincidence that the correlation between Schafer's EP habituation index and the WAIS subtests shows a rank-order correlation of +0.59 ($p < .05$) with the degree of inbreeding depression (a purely genetic effect) found on the homologous subtests of the WISC (Jensen

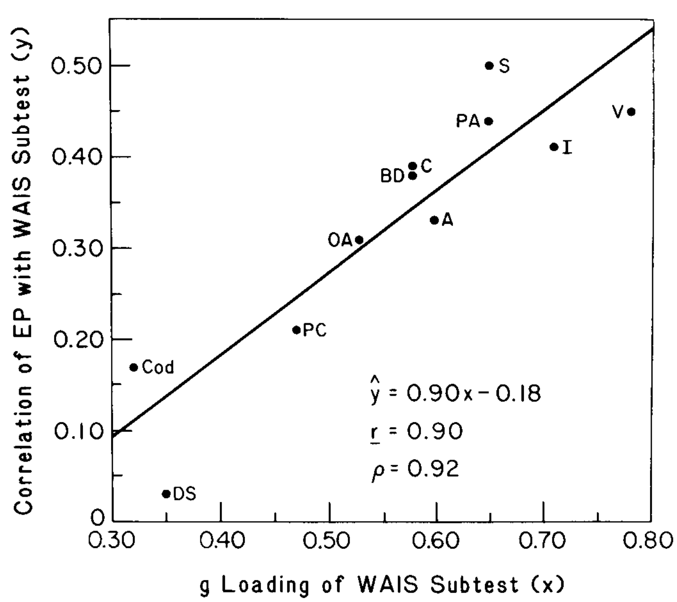


Figure 2. (Response). Correlation of the habituation index of the evoked potential (EP) with Wechsler Adult Intelligence Scale (WAIS) subtests plotted as a function of the subtests' *g* loadings, in Schafer's study. WAIS subtests: 1 – Information (I), 2 – Comprehension (C), 3 – Arithmetic (A), 4 – Similarities (S), 5 – Digit Span (DS), 6 – Vocabulary (V), 7 – Coding (Cod), 8 – Picture Completion (PC), 9 – Block Design (BD), 10 – Picture Arrangement (PA), 11 – Object Assembly (OA).

1983a). We can eagerly look forward to the working out of **Callaway**'s promising suggestions concerning the use of "psychopharmacological tools" for manipulating the biological variables underlying information processes. This biological-analytical approach is a promising avenue toward understanding the physiological substrate of *g*.

Chronometric correlates of *g*

In connection with the evoked potential studies just mentioned, it is worth noting a parallel phenomenon based on the correlation of reaction time (RT) with Wechsler subtests. P. A. Vernon (1983) extracted the general factor from a battery of elementary cognitive tasks (ECTs) in which RT was the dependent variable. The ECTs were so simple that the largest mean RTs were less than one second. The ECT general factor was substantially correlated with the WAIS Full Scale IQ, and the correlation of the general speed factor with the various WAIS subtests was related to the subtests' *g* loadings. Especially interesting is the fact that no other factors of the WAIS besides *g* showed any correlation with the ECT general speed factor. Since the target article was written, a similar recent study has come to my attention, based on the WISC-R in a sample of 59 elementary school pupils (Hemmelgarn & Kehle 1984). An apparatus very similar to that shown in Figure 8 of the target article was used. Individual differences in the slope of RT as a function of bits of information, interpreted as a measure of rate of information processing, were correlated with each of the WISC-R subtest scores (with chronological age partialed out). This profile of 12 correlations (i.e., subtests and slope of RT) showed a correlation of −0.80 ($p < .05$) with the profile of subtests' *g* loadings. The overall correlation between RT slope and WISC-R Full Scale IQ was only

−0.32 ($p < .01$); but a much higher correlation than this could hardly be expected, because it has been generally found that the slope parameter has the lowest reliability of any of the individual difference measures derived from this RT paradigm. (See Jensen, 1982a, 1982b, for detailed discussions of this RT paradigm.) Most probably, low reliability is the answer to **Carlson**'s observation that correlations between *g* and RT have not consistently shown the predicted increasing relationship across bits of information in all studies. When the means of groups differing in average IQ are used to examine slope instead of the much less reliable measures of individual differences, however, the results have been quite consistent in showing that in low-IQ groups the slope of RT across bits is greater than in high-IQ groups even when both of the contrasted groups are above the general population average in IQ.

Strategy of RT studies. There is criticism from **Carr & McDonald, Posner,** and **Rabbitt** of the fact that my presentation of correlations between RT measures and various elementary cognitive tasks (ECTs) and psychometric scores has not emphasized the same kind of analytic technique (consisting mostly of variations of Donders's subtraction method) commonly used in experimental mental chronometry. This approach is nicely summarized by Carr & McDonald. Hypothetical cognitive processes are measured indirectly by subtracting the RT for a task in which a particular process is believed to be absent from the RT for a task in which the process is believed to be required for successful performance. The remainder is a measure (usually in msec) of the time taken by the hypothesized mental process on which the two tasks are presumed to differ. I agree that this methodology is highly desirable and ultimately essential in the chronometric study of individual differences and their relation to psychometric variables. However, I considered it a highly inefficient strategy for initially exploring relationships between chronometric and psychometric variables. Those investigators who have pursued only the experimental psychology of RT, divorced from its possible relationship to individual differences in psychometric factors, may have forgotten that just a few years ago it was conventional wisdom in psychology that RT had no relationship to intelligence. Almost every psychology undergraduate has been taught in lectures and textbooks that the Galton-Cattell (i.e., James McKeen Cattell) "brass instrument" attempt to measure intelligence by means of RT and various tests of sensory discrimination was an utter failure, without learning specifically why it was a failure, and that only very complex or achievement-type tests are capable of reflecting (or defining) what psychologists mean by "intelligence." This has now been conclusively disproved by a great many recent studies. But prior to about ten years ago, I found surprising resistance to – and often scoffing rejection of –the idea that Galton and Cattell may have been right, or at least partly right, after all. It was apparent that a correlation between RT and psychometric *g* would take a lot of "proving" even for most psychologists to come to agree that there might be something worth investigating in this realm. A broad-gauged or "shotgun" search for correlations and mean differences between criterion groups selected from different sectors of the IQ distribution seemed the best strategy. Why invest a great deal of experimental refinement in some chronometric technique before establishing that at least some of the RT parameters it yields are significantly correlated with the individual difference variable of primary interest, that is, psychometric *g*, with all its obviously important scholastic, occupational, and social correlates? Whatever correlations might exist would be revealed by the raw RT measures (and such simple parameters as slope and intraindividual [trial-to-trial] variability in RT) just as well as, if not better than, the complex derived measurements of the processes hypothesized to be involved in performance on the chronometric tasks. These complex measures usually consist merely of different linear combinations of the raw RT measurements, and so any correlation that the derived measures might have with test scores would also necessarily be revealed by multiple regression analysis of the raw RT measurements. Moreover, correlational studies require good-sized samples, which, at least in exploratory research, necessitates using relatively few RT trials per subject, at the expense of achieving high reliability of individual measurements. Derived measures, being based largely on difference scores, magnify the effects of unreliability and hence further attenuate the possible correlations between RT and psychometric variables, rendering the search for correlations liable to Type II error. It is surprising that **Nettelbeck** does not seem to have noticed how seriously this very kind of Type II error has vitiated the results of the recent study by Borkowski and Krause (1983), which Nettelbeck views so uncritically. I have noted the shortcomings of this study in detail elsewhere (Jensen 1985).

Another factor in my reluctance to dive into a componential type of analysis of chronometric data in this initial exploratory stage of our research is based on what I have learned from R. J. **Sternberg**'s experience. This is the fact that there is a general RT factor (or "regression constant," as Sternberg usually terms it) in a variety of chronometric variables that is more highly correlated with psychometric *g* than most of the measurements representing specific cognitive processes (or "components," in Sternberg's terminology). In summarizing the research on the componential analysis of chronometric tasks and the correlation of components with IQ, or *g*, Sternberg and Gardner (1982) make the following observation:

> A result that at first glance appears most peculiar has emerged from many of these task analyses. . . . The regression intercept, or global "constant," often turns out to be as highly correlated or more highly correlated with scores from IQ tests than are the analyzed parameters representing separated sources of variance. Since the constant includes speed of response, e.g., button pressing, one could interpret such results trivially as indicating that motor speed is an essential ingredient of intelligence. A more plausible interpretation, and, as it will turn out, one more consistent with the bulk of the data, is that there are certain constancies in information-processing tasks that tend to be shared across wide variations in item types. We suggest that the search for the general component(s) and the search for the general factor are one and the same search–that whatever it is that leads to a unitary source of individual differences across subjects also leads to a unitary source of difference across stimulus types. (Pp. 232–33)

So before focusing on specific cognitive processes, or components, we have tried to establish firmly the correlation between the general factor of RT tasks and psychometric *g*. We are interested in whatever significant correlations we find, regardless of whether or not they are consistent with any theoretical preconceptions that we or anyone else may have had. When critics gleefully point out some theoretically unexpected effect, such as that movement time (MT) is sometimes about as highly correlated with *g* as RT, or that the RT intercept shows a higher correlation with *g* than does slope in some samples, as if they had scored a crucial point, I cannot keep from smiling. Are such findings to be put down as a loss? Theories are so tentative in this field at present that one must place more emphasis on discovering empirical relationships than on testing any specific theory. I regard any significant and replicable correlations that are unexpected in terms of general theoretical preconceptions as no less interesting than those that confirm a particular theoretical preconception. We have indeed had many surprises in our RT research so far; when they are reliable and replicable they are perfectly suitable material for theory and further inquiry. A certain "critical mass" of firmly established empirical relationships seems to be a necessary prerequisite for efficiently pursuing the kind of theory-oriented strong-inference research extolled by **Callaway,** which I agree is called for in the next phase of this program of research, now that it has been quite thoroughly demonstrated that our several chronometric paradigms yield various individual difference parameters that are indeed reliably related to psychometric *g*.

Specific criticisms of the RT research. It is always possible for critics to ignore the overall consistencies in a number of related studies and to invent ad hoc hypotheses that would seem to explain, or more usually to explain away, the results of any particular study. I am not willing to agree, however, that, because it is theoretically impossible to construct an ideally perfect lens, or because there is always some degree of atmospheric perturbation of light rays, astronomy is an altogether impossible science. The fact that it may be possible to find certain experimental paradigms, conditions, or testing procedures under which chronometric variables are not significantly correlated with psychometric variables is of no great concern, since we are seeking those conditions which *do* show correlations. And we are finding them. From our standpoint, those RT conditions which fail to yield correlations with *g* are of interest for that reason alone, but they have no theoretical refutational power whatsoever, as long as other conditions do in fact show reliable, replicable correlations with *g*.

Rabbitt surmises that the experimental separation of RT and MT in our chronometric procedures could result in a strategy, presumably adopted by the more intelligent subjects, in which there is a trade-off between RT and MT, such that subjects can shorten their RTs by responding *before* actually making a choice decision and then "hovering" to make the decision before executing the MT part of the response. **Carr & McDonald** raise essentially the same question. If this strategy were indeed in effect, we should predict a *negative* correlation between RT and MT both *within* subjects (from trial to trial) and *between* subjects (i.e., the subjects with faster RTs showing slower MTs), as well as correlations of opposite sign between *g* and RT and MT. We have long since examined all of these possibilities in our data and the results do not bear them out in the least: RT and MT are completely uncorrelated *within* subjects and *positively* correlated *between* subjects; and we have never found correlations of RT and MT with intelligence that are of opposite sign. Also of considerable interest is our finding that variation in task complexity is strongly reflected in RT but hardly at all in MT. A recent study in our laboratory, involving 14 variations in task complexity (all yielding median RTs within the range of about 600 to 1300 msec), found that the RTs on each of the tasks were much more highly correlated with Raven Matrices scores than with the MTs on the same tasks (Paul 1985). Rabbitt also conjectures that group differences in choice RT might diminish or disappear if RT trials were continued long enough for the groups to reach asymptotic levels of RT. In one study (see Jensen 1982b, p. 105) in which a group of 10 subjects was run on the Hick choice RT paradigm for a total of 540 trials spread over 9 practice sessions, there was no significant change in mean RT beyond the first session, which was the same as our standard testing procedure. We have not yet examined the effects of extended practice on the other RT tasks in the battery. The asymptotic study that Rabbitt recommends was actually done by Noble (1969), who measured RTs on 106 black and 106 white age-matched school children given 160 trials on a four-choice discrimination RT task. The groups differed significantly (whites faster), without the least indication of asymptotic convergence of the groups' mean RTs, as shown in Figure 3.

The study by Vernon and Jensen (1984) could not, of course, be reported in every detail in the target article, but the variances (or *SDs*) and correlations of the various tasks and other information that **Rabbitt** regards as important are provided in the original article. Both **Rabbitt** and **Posner** note that tasks SD2 (physically same–different words) and SA2 (synonyms–antonyms) involve verbal content, and they claim that the verbal content, rather than the tasks' intrinsic information-processing difficulty, is probably responsible for the black–white difference on these tasks. The ambiguity in interpreting this result is fully recognized by Vernon and Jensen (1984, p. 421). Other studies designed to resolve this ambiguity are already in progress. It will be surprising to me if Posner's conjecture that differential reading skill of blacks and whites, independent of *g*, would account for the black–white difference on tasks SD2 and SA2. One statistical test would be to regress out that part of the variance in reading skill which is independent of *g* (assessed by nonverbal tests) from the RT variables and see whether a significant black–white difference in mean RTs remains. Other research indicates that when *g* is regressed out of scores on verbal tests, the black–white difference virtually disappears. That is, the difference in reading skill seems largely to reflect the more general black–white difference in *g*.

Experimental chronometricians (**Nettelbeck, Poortinga, Posner, Rabbitt**) are concerned with the phenomenon known as "speed–accuracy trade-off," suggesting that perhaps the brighter subjects adopt a strategy of sacrificing accuracy for speed, thereby showing faster RT and a higher error rate. But this trade-off seems to be mainly a *within*-subjects phenomenon, accounting for

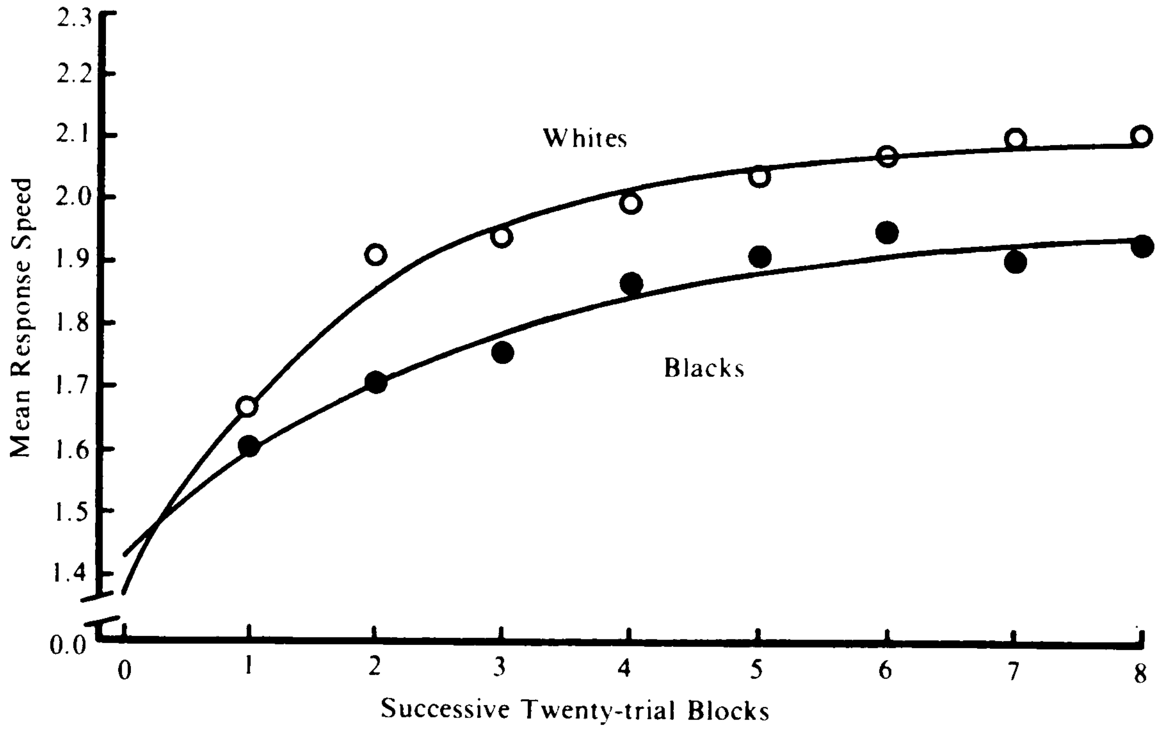


Figure 3. (Response). Mean response speed (reciprocal of RT) in successive 20-trial blocks on a 4-choice RT test. Each curve based on 106 children. (From Noble 1969.)

negative correlations (*within* subjects) between RTs and error rates under different levels of task difficulty. It has not been a problem at all in the interpretation of the correlation between individual differences in RT and g, because the *between*-subjects correlation of RT and error rate is a *positive* correlation, and both RT and error rate are negatively correlated with g. That is, the brighter subjects are both faster and more accurate than the less bright subjects; we have never found any evidence of a speed–accuracy trade-off *between* subjects in our analyses of RT data. These relationships can perhaps be seen more clearly as depicted in Figure 4. On the *simple task*, hypothetical persons A, B, and C are shown to have the same short RT and low error rate. On the *complex task*, the latent ability differences between A, B, and C are manifested as variation in their RTs and error rates. Their performances, as reflected jointly by RT and errors, will tend to fall somewhere on each of the arcs that describe the speed–accuracy trade-off and are different for each person. If the same low error rate of the simple task is to be maintained for the complex task, the RT is greatly increased for all persons (vertical line = zero speed–accuracy trade-off). If the RT in the simple task is to be maintained in the complex task, the error rate is greatly increased for all persons (horizontal line = 100% speed–accuracy trade-off). So the arc for each person describes an *inverse* relationship (or *negative* correlation) between RT and error rate. But *between* persons, RT and error rate show a *direct* relationship (or *positive* correlation). The line marked *X* in Figure 4 indicates a fairly high speed–accuracy trade-off for a typical RT study, if the error rate (on the abscissa) is assumed to range between *zero* and *chance*. Thus the shaded area represents the most desirable region for performance when studying individual differences in RT in that it spreads out individual differences in RT much more than in error rate, a feature observed in all of our RT studies. Hence the observed correlation between RT variables and g can in no way be accounted for in terms of speed–accuracy trade-off.

Jones complains that the figures showing mean group differences on the various chronometric tasks express the differences directly in terms of milliseconds, rather than in standard deviation (σ) units. I had used the raw RT

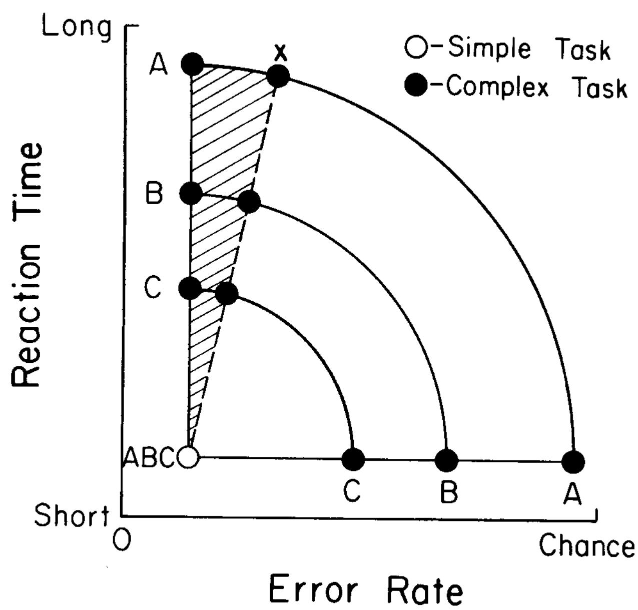


Figure 4. (Response). The idealized relationship between RT and error rate for simple and complex tasks. The arcs describe the speed–accuracy trade-off for hypothetical persons A, B, and C, who are shown here as performing equally on the simple task. Shaded area represents most desirable region of speed–accuracy trade-off for RT studies.

differences to take advantage of a luxury that is generally denied for ordinary psychometric tests, namely, a true ratio scale, which RT represents, so that the mean group differences in RT are differences in real time units, with equal intervals and a true zero point. The results depicted in Figures 11 and 12, it turns out, remain essentially the same when differences are expressed in σ units. In Figure 12, for example, when the group differences (vocational college versus university) on the tasks, expressed in σ units, are plotted as a function of task complexity as indexed by mean RT, the Pearson correlation is $+0.92$ ($\rho = +0.93$), as compared with $+0.97$ when the RT differences are expressed in msec. If instead of differences we use the ratio of vocational college/university RTs, the correlation is $+0.89$ ($\rho = +0.95$); and if the RTs are subjected to a logarithmic transformation (which tends to make the standard deviations and means uncorrelated), the corresponding correlation becomes $+0.93$ ($\rho = +0.95$). In other words, no matter what the scale is on which the group differences are expressed, the group differences are found to increase as a function of task difficulty or complexity. (The same thing is true of Figure 11.) [I am grateful for Jones's noting the errors in the target article's Figure 10, which have been duly corrected in the published version.] The other questions raised by Jones about this study are answered in the original Vernon and Jensen (1984) article.

Poortinga believes that cultural factors may affect RT in ECTs and that such tasks as simple RT may be culturally biased and hence "nonequivalent" across different populations. But the lack of evidence for cultural bias with respect to the American black and white populations in much more complex and culture-loaded psychometric tests makes it an improbable hypothesis that cultural bias would be significantly implicated in ECTs. Cultural bias could be investigated by much the same methods as have been applied to conventional tests (Jensen 1980a). Poortinga infers bias on the basis of theoretical preconceptions of the pattern of group differences one should expect for various RT parameters. This puts too much faith in the present theories of RT and ECTs. For the time being, I would avoid theoretical preconceptions about which parameters should be most meaningful and take a more direct empirical approach. This would consist of looking at differences in RT parameters between different population samples that are hypothesized to differ culturally in ways that affect performance in ECTs and comparing the pattern of differences with the corresponding patterns found in pairs of groups that are selected to be high and low in psychometric *g* but are culturally equivalent. Ideally, one could use groups of full siblings reared together, with one member of each sib pair assigned to the low-*g* group and the other member assigned to the high-*g* group. These two comparison groups would be as culturally equivalent as possible. If the two supposedly culturally different population samples show essentially the same pattern of RT differences on a number of ECTs as the culturally equivalent groups that were selected to differ in *g*, then we would be forced either to reject the cultural bias hypothesis *or* to hypothesize that the cultural difference perfectly mimics the *g* difference between two culturally equivalent groups. With enough different ECTs, the latter hypothesis becomes highly implausible. I would like to see this type of study performed with the set of RT tasks that were used in Poortinga's (1971) own interesting study.

RT and athletic skill. The black–white differences in response latencies on some of the elementary cognitive tasks is called into question by **Das** and **Whimbey** on the ground that a relatively large proportion of topnotch athletes and Olympic gold medalists are black. First, it is a mistake to try to explain a given phenomenon (black–white differences in RT) in terms of another even more complex and less well understood phenomenon (athletic skill). And a phenomenon observed in one realm (the athletic field) certainly cannot refute a questionably related phenomenon observed in another realm (the psychological laboratory). Second, the exceptional Olympic-level athletes are highly selected from their respective populations, and their particular talents may represent other features of the population distribution of ability than the central tendency, such as the variance, which would affect the remote tails of the distribution from which exceptionally talented individuals are selected. Third, the argument presumes that the order of RTs (in the range of about 200 to 1200 msec) represented in our studies constitutes a sizable proportion of the variance in athletic skills. This is most unlikely. RT evidently has much more to do with *g* than with athletic prowess. Noble (1978) lists a large number of physical fitness and body build factors, independent of psychomotor and perceptual factors, that are involved in varying degrees in different athletic skills, which generally require sequential integration of numerous separate movements of large muscle groups, whole-body coordination, and the like. It may seem even more surprising to Das and Whimbey that blacks have been found to perform significantly less well than whites even on the pursuit rotor, a simple motor learning task (Noble 1978, pp. 346–47; Payne & Turkat 1982). Apparently, very fast RT is not necessary for becoming the greatest boxer of all time. According to Keele (1973, as cited by Hunt 1976, p. 238), "Muhammad Ali, a heavyweight boxer who, in his prime, was lauded for his 'cat-like reflexes,' had a quite average motor reaction time."

The genetic heritability issue

Several commentators (**Bardis, Cattell, Johnson & Nagoshi,** and **Stanovich**) bring up the genetic question. However, I have consistently treated Spearman's hypothesis as a *phenotypic* phenomenon. Strictly speaking, neither the data nor the methodology of the target article permits inferences about the relative roles of genetic and nongenetic sources of variance in the observed, or phenotypic, population differences. Stanovich is perfectly right in noting that the findings are moot regarding the *causes* of the differences. I have long since concluded that the only technically available method, at present, that would permit proper genetic inferences regarding population differences in IQ (or in any other phenotype) would be to perform a true genetic experiment, cross-mating random samples of the two populations and cross-fostering the offspring. But socially and ethically such an experiment would be wholly unfeasible and impermissible. All other feasible lines of research can at most only diminish or augment the subjective plausibility of the hypothesis that genetic factors are involved in any particular physical or mental trait difference between populations. The broad evolutionary context of biological and behavioral variables in which **Rushton** finds remarkably systematic relationships among differences between populations of African, Asian, and European origin affords a much needed perspective for further advances in the study of human variation, although such research will unfortunately invite still more controversy and even opprobrium in the ideological climate that currently prevails in the social sciences.

Individual variation *within* populations is quite another matter, however. It is now well established that genetic factors are strongly involved in individual differences on psychometric tests. (**Bardis** is simply wrong on this issue, and he errs in believing that the estimation of heritability depends on the direct measurement of environmental factors.) But ECTs have not yet been subjected to extensive genetic analysis. The only published genetic study of ECTs that I am aware of is based on several ECTs quite similar to those described in the target article, administered to a total of 47 pairs of monozygotic and dizygotic twins reared apart, from which the authors (McGue, Bouchard, Lykken & Feuer 1984) concluded:

> The results reported here support the existence of a general speed component underlying performance on most experimental cognitive tasks which is strongly related to psychometric measures of "g," and for which there are substantial genetic effects. Although much of the relationship between psychometric test performance and processing speed may be attributed to the relationship between this general speed factor and "g," we did find evidence for a second component which loads on measures of the rate of specific cognitive processes, which was specifically associated with psychometric measures of verbal ability, and which appeared to have little or no genetic basis. (P. 256)

The social context of *g*

The only commentator who brings Spearman's hypothesis directly and specifically into apposition with its real-life social and economic consequences, is **Cattell,** who predicts that the percentage of blacks in different occupations should be inversely related to the mean intelligence levels of persons employed in the occupations. If shown to be true, this prediction would mean, of course, that disparities in the proportional representation of black and white workers in various occupational categories are not mainly attributable to prejudice and discrimination in hiring, but are due to differences in measurable *g*-loaded abilities, whatever the cause of the differences. I have not looked into data on this point myself, but quite precise data on a range of occupations (ranging from physician and engineer to truck driver and meat cutter), directly aimed at Cattell's prediction, have been assembled by Linda Gottfredson (personal communication), a sociologist at the Johns Hopkins University. In light of Cattell's query, it would be most valuable if Gottfredson submitted this analysis to Continuing Commentary. Gottfredson's analysis, based on 1970 and 1980 statistics from the U.S. Department of Labor and the Bureau of the Census, strikingly bears out Cattell's prediction, with a near perfect rank-order correlation between the theoretically expected and the observed ratios of black to white employees in different occupations.

I suppose it is largely because of my investigating phenomena such as Spearman's hypothesis, which have such crucial and sensitive social correlates, that perhaps quite a few psychologists share in **Sternberg**'s emotional "distaste" for my study of black–white differences (also voiced in different tones by **Bardis** and **Das**). I make no apology for my choice of research topics. I think that my own nominal fields of expertise (educational and differential psychology) would be remiss if they shunned efforts to describe and understand more accurately one of the most perplexing and critical of current problems. Of all the myriad subjects being investigated in the behavioral and social sciences, it seems to me that one of the most easily justified is the black–white statistical disparity in cognitive abilities, with its far-reaching educational, economic, and social consequences. Should we not apply the tools of our science to such socially important issues as best we can? The success of such efforts will demonstrate that psychology can actually behave as a science in dealing with socially sensitive issues, rather than merely rationalize popular prejudice and social ideology.

References

Ahern, S. & Beatty, J. (1979) Pupillary responses during information processing vary with Scholastic Aptitude Test scores. *Science* 205:1289–92. [rARJ]

Allport, D. A. (1980) Patterns and actions: Cognitive mechanisms are content-specific. In: *Cognitive psychology: New directions,* ed. G. Claxton. Routledge & Kegan Paul. [THC]

Anderson, R. J. & Sisco, F. H. (1977) *Standardization of the WISC-R performance scale for deaf children.* Office of Demographic Studies Publication Series T, no. 1. Gallaudet College. [taARJ]

Ashcraft, M. H. & Stazyk, E. H. (1981) Mental addition: A test of three verification models. *Memory & Cognition* 9:185–96. [THC]

Baddeley, A. D. & Hitch, G. (1974) Working memory. In: *The psychology of learning and motivation,* vol. 8, ed. G. Bower. Academic Press. [PMAR]

Baker, J. R. (1974) *Race.* Oxford University Press. [JPR]

Bardis, P. D. (1969) The principle of instrumental parsimony. *Revue Internationale de Sociologie:* 92–101. [PDB]

Baron, J. (in press) *Rationality and intelligence.* Cambridge University Press. [JB]

Baron, J. & Treiman, R. (1980) Some problems in the study of differences in cognitive processes. *Memory and Cognition* 8:313–21. [JB]

Bengtson, V. L., Kasschau, P. L. & Ragan, P. K. (1977) The impact of social structure on aging individuals. In: *Handbook of the psychology of aging,* ed. J. E. Birren & K. W. Schaie. Van Nostrand Reinhold. [JPR]

Bethge, H. J., Carlson, J. S. & Wiedl, K. H. (1982) The effects of dynamic assessment procedures on Raven matrices performance, visual search behavior, test anxiety and test orientation. *Intelligence* 6:89–97. [JSC]

Blanchard, J. (1980) Preliminary investigation of transfer between single-word decoding ability and contextual reading comprehension by poor readers in grade six. *Perceptual and Motor Skills* 51:1271–81. [KES]

Blau, Z. (1981) *Black children/white children.* Free Press. [KES]

Bock, R. D. & Mislevy, R. J. (1981) *The profile of American youth: Data quality analysis of the Armed Services Vocational Aptitude Battery.* University of Chicago/National Opinion Research Center. [taARJ]

Borkowski, J. G. (1985) Signs of intelligence: Strategy generalization and metacognition. In: *The growth of reflection in children,* ed. S. Yussen. Academic Press. [JGB]

Borkowski, J. G. & Krause, A. (1983) Racial differences in intelligence: The importance of the executive system. *Intelligence* 1:379–95. [JGB, rARJ, TN, KES]

Borkowski, J. G. & Peck, V. A. (in press) Causes and consequences of metamemory in gifted children. In: *Conceptions of intelligence,* ed. R. Sternberg & J. Davidson. Cambridge University Press. [JGB]

Bouchard, T. J., Jr. (1982) Twins: Nurture's twice-told tale. In: *1983 Yearbook of science and the future*. Encylopaedia Britannica. [JPR]

Bouchard, T. J., Jr. & McGue, M. (1981) Familial studies of intelligence: A review. *Science* 212:1055–59. [JPR]

Braden, J. P. (1984) The factorial similarity of the WISC-R Performance Scale in deaf and hearing samples. *Personality and Individual Differences* 5:403–9. [taARJ]

Bradley, L. & Bryant, P. (1983) Categorizing sounds and learning to read–A causal connection. *Nature* 301:419–21. [KES]

Bradley, R. & Caldwell, B. (1984) The relation of infants' home environments to achievement test performance in first grade: A follow-up study. *Child Development* 55:803–9. [KES]

Brand, C. R. (1984) Personality dimensions: An overview of modern trait psychology. In: *Psychological Survey 5*, ed. J. Nicholson & H. Beloff. British Psychological Society. [CB]

(1985a) The psychological bases of political attitudes and interests. In: *Hans Eysenck: Consensus and controversy*, ed. S. Modgil & C. Modgil. Falmer. [CB]

(1985b) Intelligence and inspection time: An ontogenetic relationship? In: *The biology of human intelligence*, ed. C. Turner. Eugenics Society. [CB]

Bridgeman, B. & Buttram, J. (1975) Race differences in nonverbal analogy test performance as a function of verbal strategy training. *Journal of Educational Psychology* 67:586–90. [JSC]

Bulmer, M. G. (1970) *The biology of twinning in man*. Clarendon Press. [JPR]

Butterfield, E. C. & Ferretti, R. P. (in press) Toward a theoretical integration of cognitive hypotheses about intellectual differences among children. In: *Intelligence and cognition in special children*, ed. J. Borkowski & J. Day. Ablex. [JGB]

Callaway, E. (1975) *Brain electrical potentials and individual psychological differences*. Grune & Stratton. [EC, taARJ]

(1984) Human information processing: Some effects of methylphenidate, age and scopolamine. *Biological Psychiatry* 19:649–62. [EC]

Callaway, E., Halliday, R., Naylor, H. & Schechter, G. (in press) Effects of oral scopolamine on human stimulus evaluation. *Psychopharmacology*. [EC]

Carlson, J. S. (1983) Applications of dynamic assessment to cognitive and perceptual functioning of three ethnic groups. National Institute of Education final report, grant #NIE-G-81-0081, 100 pp. [JSC]

Carlson, J. S. & Jensen, C. M. (1982) Reaction time, movement time, and intelligence: A replication and extension. *Intelligence* 6:265–74. [JGB, taARJ, TN]

Carlson, J. S., Jensen, C. M. & Widaman, K. F. (1983) Reaction time, intelligence and attention. *Intelligence* 7:329–44. [JSC]

Carlson, J. S. & Wiedl, K. H. (1980) Applications of a dynamic testing approach in intelligence assessment: Empirical results and theoretical formulations. *Zeitschrift für Differentielle Psychologie* 1:303–18. [JSC]

Carr, T. H. (1981) Building theories of reading ability: On the relation between individual differences in cognitive skills and reading ability. *Cognition* 9:73–114. [THC]

(1984) Attention, skill, and intelligence: Some speculations on extreme individual differences in human performance. In: *Learning and cognition in the mentally retarded*, ed. P. H. Brooks, R. Sperber & C. McCauley. Erlbaum. [THC]

Carr, T. H., Brown, T. L. & Vavrus, L. G. (1985) Using component skills analysis to integrate findings about reading development. In: *New directions in child development #27: The development of reading skills*, ed. T. H. Carr. Jossey-Bass. [THC]

Carr, T. H., Pollatsek, A. & Posner, M. I. (1981) What does the visual system know about words? *Perception and Psychophysics* 29:183–90. [THC]

Carr, T. H., Posner, M. I., Pollatsek, A. & Snyder, C. R. R. (1979) Orthography and familiarity effects in word processing. *Journal of Experimental Psychology: General* 108:389–414. [MIP]

Carroll, J. B. (1980) Individual difference relations in psychometric and experimental cognitive tasks. Report no. 163, L. L. Thurstone Psychometric Laboratory, University of North Carolina. [PK]

(1981) Review of *Bias in mental testing*, A. R. Jensen. *Psychometrika* 46:227–33. [LVJ]

Cattell, R. B. (1940) A culture-fair intelligence test 1. *Journal of Educational Psychology* 31:161–79. [RBC]

(1963) Theory of fluid and crystalized intelligence: A critical experiment. *Journal of Educational Psychology* 54:1–22. [RBC, J-EG, taARJ]

(1967) The theory of fluid and crystallized intelligence checked at the 5-year-old level. *British Journal of Educational Psychology* 37:209–34. [RBC]

(1971) *Abilities: Their structure, growth and action*. IPAT, Champaign, Ill. [RBC]

(1978) *The scientific use of factor analysis in behavioral and life sciences*. Plenum. [taARJ]

Chase, W. G. & Ericsson, K. A. (1981) Skilled memory. In: *Cognitive skills and their acquisition*, ed. J. R. Anderson. Erlbaum. [PMAR]

Chase, W. G. & Simon, H. A. (1973) Perception in chess. *Cognitive Psychology* 4:55–81. [THC]

Chi, M. T. H. (1976) Short-term memory limitations in children: Capacity or processing deficits? *Memory and Cognition* 4:559–72. [MIP]

Cole, M., Gay, G. & Glick, J. (1968) Some experimental studies of Kpelle quantitative behavior. *Psychonomic Monograph Supplement* 2, no. 10 (Whole no. 26):173–90. [YHP]

Coltheart, M. (1981) Disorders of reading and their implications for models of normal reading. *Visible Language* 15:245–86. [MIP]

Cooper, L. & Regan, D. T. (1982) Attention, perception and intelligence. In: *Handbook of human intelligence*, ed. R. Sternberg. Wiley. [PMAR]

Cronbach, L. J. (1957) The two disciplines of scientific psychology. *American Psychologist* 12:671–84. [CB]

(1969) Heredity, environment and educational policy. *Harvard Educational Review* 39:338–47. [JPD]

(1979) The Armed Services Vocational Aptitude Battery–A test battery in transition. *Personnel and Guidance Journal* 57:232–37. [taARJ]

Dasen, P. R. (1984) The cross-cultural study of intelligence: Piaget and the Baoulé. *International Journal of Psychology*. In press. [JSC]

Day, J. E., French, L. & Hall, L. K. (1985) Social interaction and cognitive development. In: *Metacognition, cognition, and human performance*, ed. D. L. Forrest-Pressley, G. E. MacKinnon & T. G. Waller. Academic Press. [JGB]

Dempster, F. N. (1981) Memory span: Sources of individual and developmental differences. *Psychological Bulletin* 89:63–100. [CB]

(1985) Short-term memory development in childhood and adolescence. In: *Basic processes in memory development: Progress in cognitive development research*, ed. C. J. Brainerd & M. Pressley. Springer. [CB]

Detterman, D. (1982) Does "g" exist? *Intelligence* 6:99–108. [KES]

Detterman, D. K. & Sternberg, R. J., eds. (1982) *How and how much can intelligence be increased*. Ablex. [taARJ]

Dillon, R. & Carlson, J. S. (1978) The use of activation variables in the assessment of cognitive abilities in three ethnic groups: A testing-the-limits approach. *Educational and Psychological Measurement* 38:437–43. [JSC]

Donders, F. C. (1868/69; 1969) Over de snelheid van psychische processen. Trans. W. G. Koster. In: *Attention and performance II*, ed. W. G. Koster. *Acta Psychologica* 30:412–31. [THC]

Duncan Johnson, C. (1981) P300 latency: A new metric of information processing. *Psychophysiology* 18:207–15. [EC]

Durkin, D. (1982) A study of poor black children who are successful readers. Reading Education Report no. 33. Center for the Study of Reading, University of Illinois. [KES]

Engle, R. W. & Bukstel, L. (1978) Memory processes among bridge players of different expertise. *American Journal of Psychology* 91:673–89. [THC]

Ertl, J. P. (1969) Neural efficiency and human intelligence. Final report, U.S. Office of Education Project no. 9-0105. [EC, RBC]

Ertl, J. P. & Schafer, E. W. P. (1967) Cortical activity preceding speech. *Life Sciences* 6:473–79. [EC]

Evans, M. A. & Carr, T. H. (1985) Cognitive abilities, conditions of learning, and the early development of reading skill. *Reading Research Quarterly*. In press. [THC]

Eveleth, P. B & Tanner, J. M. (1976) *Worldwide variation in human growth* Cambridge University Press. [JPR]

Eysenck, H. J. (1939) Review of *Primary mental abilities* by L. L. Thurstone. *British Journal of Educational Psychology* 9:270–75. [taARJ]

(1979) *The structure and measurement of intelligence*. Springer. [HJE]

(1982a) The psychophysiology of intelligence. In: *Advances in personality assessment*, vol. 1, ed. C. D. Spielberger & J. N. Butcher. Erlbaum. [RBC, taARJ]

(1982b) A *model of intelligence*. Springer. [PEV]

(1984) The effect of race on human abilities and mental test scores. In: *Perspectives on bias in mental testing*, ed. C. R. Reynolds & R. T. Brown. Plenum. [HJE]

(1985) The theory of intelligence and the psychophysiology of cognition. In: *Advances in the psychology of human intelligence*, vol. 3, ed. R. J. Sternberg. Erlbaum. [HJE]

Eysenck, H. J. & Barrett, P. (1985) Psychophysiology and the measurement of intelligence. In: *Methodological and statistical advances in the study of*

individual differences, ed. C. R. Reynolds & V. Willson. Plenum. [rARJ, EWPS]
Feuerstein, R. (1979) *The dynamic assessment of retarded performers.* University Park Press. [KES]
Fitts, P. M. & Posner, M. I. (1967) *Human performance.* Wadsworth. [YHP]
Floderus-Myrhed, B., Pedersen, N. & Rasmuson, I. (1980) Assessment of heritability for personality, based on a short form of the Eysenck Personality Inventory: A study of 12,898 twin pairs. *Behavior Genetics* 10:153–62. [JPR]
Flynn, J. (1980) *Race, IQ, and Jensen.* Routledge & Kegan Paul. [KES]
(1984) Japanese IQ. *Nature* 308:222. [JPR]
Fraser, I. C. (1984) The psychophysiological measurement of adult intelligence. Thesis, Department of Psychology, University of Edinburgh. [CB]
Frederiksen, J. R. (1980) Component skills in reading: Measurement of individual differences through chronometric analysis. In: *Aptitude, learning, and instruction*, Vol. 1: *Cognitive process analyses of aptitude*, ed. R. E. Snow, P.-A. Federico & W. E. Montague. Erlbaum. [THC]
Freedman, D. G. (1979) *Human sociobiology: A holistic approach.* Free Press. [JPR]
French, J. W., Ekstrom, R. B. & Price, L. A. (1963) *Kit of reference tests for cognitive factors.* Educational Testing Service. [rARJ]
Garrett, H. E., Bryan, A. I. & Perl, R. E. (1935) The age factor in mental organization. *Archives of Psychology*, no. 176. [rARJ]
Gordon, R. A. (1976) Prevalence: The rare datum in delinquency measurement and its implications for the theory of delinquency. In: *The juvenile justice system*, ed. M. W. Klein. Sage. [RAG]
(1984) Digits backward and the Mercer-Kamin law: An empirical response to Mercer's treatment of internal validity of IQ tests. In: *Perspectives on bias in mental testing*, ed. C. R. Reynolds & R. T. Brown. Plenum. [RAG]
Gorsuch, R. L. (1974) *Factor analysis.* W. B. Saunders. [RAG, rARJ]
Gray, J. A. (1982) *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system.* Oxford University Press. [JPR]
Groen, G. J. & Parkman, J. M. (1972) A chronometric analysis of simple addition. *Psychological Review* 79:329–43. [THC]
Guilford, J. P. (1965) *Fundamental statistics in psychology and education.* 4th ed. McGraw-Hill. [RAG]
Gustafsson, J.-E. (1984) A unifying model for the structure of intellectual abilities. *Intelligence* 8:179–203. [J-EG, rARJ]
Guthrie, J. T. (1973) Models of reading and reading disability. *Journal of Educational Psychology* 65:9–18. [THC]
Haier, R. J., Robinson, D. L., Braden, W. & Williams, D. (1983) Electrical potentials of the cerebral cortex and psychometric intelligence. *Personality and Individual Differences* 4:591–99. [rARJ]
Hansen, J. & Pearson, P. (1983) An instructional study: Improving the inferential comprehension of good and poor fourth-grade readers. *Journal of Educational Psychology* 75:821–29. [KES]
Harlow, H. F. (1959) Learning set and error factor theory. In: *Psychology: A study of a science*, vol. 2, ed. S. Koch. McGraw-Hill. [rARJ]
Harman, H. H. (1960) *Modern factor analysis.* University of Chicago Press. [RAG]
(1967) *Modern factor analysis.* 2nd ed. University of Chicago Press. [taARJ]
Hebb, D. O. (1949) *The organization of behavior.* Wiley. [PEV]
Heber, R. & Garber, H. (1973) The Milwaukee Project: A study of the use of family investigation to prevent cultural–familial mental retardation. In: *Exceptional infant*, vol. 3, ed. B. Z. Friedlander, G. M. Sterritt & G. K. Kirk. Brunner/Mazel. [JSC]
Hemmelgarn, T. E. & Kehle, T. J. (1984) The relationship between reaction time and intelligence in children. *School Psychology International* 5:77–84. [rARJ]
Hendrickson, A. E. (1982) The biological basis of intelligence, Part 1: Theory. In: *A model for intelligence*, ed. H. J. Eysenck. Springer. [CB]
Hendrickson, D. E. (1982) The biological basis of intelligence, Part 2: Measurement. In: *A model for intelligence*, ed. H. J. Eysenck. Springer. [CB]
Hendrickson, D. E. & Hendrickson, A. E. (1980) The biological basis of individual differences in intelligence. *Personality and Individual Differences* 1:3–33. [taARJ, EWPS]
Hennessy, J. J. (1974) *Structure and patterns of mental abilities in several ethnic groups.* Doctoral dissertation, New York University. University Microfilms no. 74–24, 997. [taARJ]
Hennessy, J. J. & Merrifield, P. R. (1976) A comparison of the factor structures of mental abilities in four ethnic groups. *Journal of Educational Psychology* 68:754–59. [RAG, taARJ]
Hogan, D. D. (1971) Cortical response of retardates for AER and audiometry. *American Journal of Mental Deficiency* 75:474. [EC]
Horn, J. L. (1965) Fluid and crystallized intelligence, a factor analytic study of the structure among primary mental abilities. Doctoral dissertation, University of Illinois. [RBC]
(1966) Short period fluctuations in intelligence. University of Denver Research Institute. [RBC]
(1968) Organization of abilities and the development of intelligence. *Psychological Review* 75:242–59. [RBC, J-EG]
(in press) Remodeling old models of intelligence. In: *Handbook of intelligence*, ed. B. Wolman. Prentice-Hall. [JGB]
Hulme, C. (1984) *British Journal of Psychology.* In press. [PMAR]
Humphreys, L. G. (1971) Theory of intelligence. In: *Intelligence: Genetic and environmental influences*, ed. R. Cancro. Grune & Stratton. [rARJ, LVJ]
(1981) The primary mental ability. In: *Intelligence and learning*, ed. M. P. Friedman, J. P. Das & N. O'Connor. Plenum. [taARJ]
(1983) Review of *Ability testing* by A. K. Wigdor & W. R. Garner. *American Scientist* 71:302–3. [taARJ]
(1984) General intelligence. In: *Perspectives on bias in mental testing*, ed. C. R. Reynolds & R. T. Brown. Plenum. [LVJ]
Humphreys, L. G., Fleishman, A. I. & Lin, P.-C. (1977) Causes of racial and socioeconomic differences in cognitive tests. *Journal of Research in Personality* 11:191–208. [taARJ]
Hunt, E. (1976) Varieties of cognitive power. In: *The nature of intelligence*, ed. L. B. Resnick. Erlbaum. [rARJ]
(1978) Mechanics of verbal ability. *Psychological Review* 85:109–30. [PMAR]
(1980) Intelligence as an information-processing concept. *British Journal of Psychology* 71:449–74. [EC]
Hunt, E., Lunneborg, C. & Lewis, J. (1975) What does it mean to be high verbal? *Cognitive Psychology* 7:194–227. [PMAR]
Hunt, J. (1980) *Early development and experience*, vol. 10. 1974 Heinz Werner Lecture Series. [PDB]
Hunt, J. McV. (1961) *Intelligence and experience.* Ronald. [rARJ]
Hunter, J., Jones, L., Vincent, H. & Carmichael, J. W. (1982) Project SOAR: Teaching cognitive skills in a pre-college program. *Journal of Learning Skills* 2:24–26. [AW]
IPAT. (1950, 1959) The Culture Fair Intelligence Test Series. IPAT, Champaign, Ill. [RBC]
Jackson, M. D. & McClelland, J. L. (1979) Processing determinants of reading speed. *Journal of Experimental Psychology: General* 108:151–81. [THC]
Jarman, R. F. (1978) Level I and Level II abilities: Some theoretical reinterpretations. *British Journal of Psychology* 69:257–69. [JPD]
Jensen, A. R. (1966) Social class and perceptual learning. *Mental Hygiene* 50:226–39. [PHS]
(1969) How much can we boost IQ and scholastic achievement? *Harvard Educational Review* 39:1–123. [JPD]
(1971) Do schools cheat minority children? *Educational Research* 14:3–28. [rARJ]
(1973a) *Educability and group differences.* Methuen/Harper & Row. [taARJ, RCN]
(1973b) Level I and Level II abilities in three ethnic groups. *American Educational Research Journal* 4:263–76. [tarARJ]
(1974a) Interaction of Level I and Level II abilities with race and socioeconomic status. *Journal of Educational Psychology* 66:99–111. [tarARJ]
(1974b) How biased are culture-loaded tests? *Genetic Psychology Monographs* 90:185–244. [taARJ]
(1977a) An examination of culture bias in the Wonderlic Personnel Test. *Intelligence* 1:51–64. [taARJ]
(1977b) Cumulative deficit in IQ of blacks in the rural south. *Developmental Psychology* 13:184–91. [PDB, KES]
(1979) *g*: Outmoded theory or unconquered frontier? *Creative Science and Technology* 2:16–29. [taARJ]
(1980a) *Bias in mental testing.* Free Press. [JGB, tarARJ, LVJ, RCN, YHP, PHS]
(1980b) Chronometric analysis of intelligence. *Journal of Social and Biological Structures* 3:103–22. [taARJ, TN]
(1981) Reaction time and intelligence. In: *Intelligence and learning*, ed. M. P. Friedman, J. P. Das & N. O'Connor. Plenum. [taARJ, PMAR]
(1982a) The chronometry of intelligence. In: *Advances in the psychology of human intelligence*, vol. 1, ed. R. J. Sternberg. Erlbaum. [tarARJ, PMAR]
(1982b) Reaction time and psychometric *g*. In: *A model for intelligence*, ed. H. J. Eysenck. Springer. [JB, tarARJ, YHP]
(1983a) The effects of inbreeding on mental ability factors. *Personality and Individual Differences* 4:71–87. [tarARJ, RCJ]

(1983b) The definition of intelligence and factor score indeterminacy. *Behavioral and Brain Sciences* 6:313–15. [taARJ]
(1983c) Again, how much can we boost IQ? Review of *How and how much can intelligence be increased*, ed. D. K. Detterman & R. J. Sternberg. *Contemporary Psychology* 28:756:58. [taARJ]
(1984a) Test validity: *g* versus the specificity doctrine. *Journal of Social and Biological Structures* 7:93–118. [taARJ, AW]
(1984b) The black–white difference on the K-ABC: Implications for future tests. *Journal of Special Education* 18:377–408. [taARJ]
(1984c) Test bias: Concepts and criticisms. In: *Perspectives on bias in mental testing*, ed. C. R. Reynolds & R. T. Brown. Plenum. [rARJ]
(1984d) Sociobiology and differential psychology: The arduous climb from plausibility to proof. In: *Annals of theoretical psychology*, vol. 2, ed. J. R. Royce & L. P. Mos. Plenum. [JPR]
(1985) Race differences and Type II errors: A comment on Borkowski and Krause. *Intelligence* 9:33–39. [rARJ]
(in press) Methodological and statistical techniques for the chronometric study of mental abilities. In: *Methodological and statistical advances in the study of individual differences*, ed. C. R. Reynolds & V. L. Willson. Plenum. [taARJ]
Jensen, A. R. & Figueroa, R. A. (1975) Forward and backward digit span interaction with race and IQ: Predictions from Jensen's theory. *Journal of Educational Psychology* 67:882–93. [tarARJ]
Jensen, A. R. & Inouye, A. R. (1980) Level I and Level II abilities in Asian, white, and black children. *Intelligence* 4:41–49. [rARJ]
Jensen, A. R. & Munro, E. (1979) Reaction time, movement time and intelligence. *Intelligence* 3:121–26. [taARJ]
Jensen, A. R. & Osborne, R. T. (1979) Forward and backward digit span interaction with race and IQ: A longitudinal developmental comparison. *Indian Journal of Psychology* 54:75–87. [taARJ]
Jensen, A. R. & Reynolds, C. R. (1982) Race, social class and ability differences on the WISC-R. *Personality and Individual Differences* 3:423–38. [RAG, tarARJ, KES]
Jensen, A. R., Schafer, E. W. P. & Crinella, F. M. (1981) Reaction time, evoked brain potentials and psychometric *g* in the severely retarded. *Intelligence* 5:179–97. [taARJ]
Jerison, H. J. (1973) *Evolution of brain and intelligence*. Academic Press. [rARJ]
Jones, L. V. (1984) White–black achievement differences: The narrowing gap. *American Psychologist* 39. In press. [LVJ]
Karrer, R. (1984) An analysis of input, central, and motor segments of response time in the mentally retarded. Paper presented at the Conference on Motor Behavior of the Retarded, NICHD, Washington, D.C. [JPD]
Kaufman, A. S. & Kaufman, N. L. (1983) *Kaufman Assessment Battery for Children: Interpretive manual*. American Guidance Service. [RAG, taARJ]
Keele, S. W. (1973) *Attention and human performance*. Goodyear. [rARJ]
Kendall, M. G. & Stuart, A. (1976) *The advanced theory of statistics*, vol. 2, *Inference and relationship*. 3rd ed. Griffin. [taARJ]
Kinchla, R. A. (1974) Detecting target elements in multi-element arrays: A confusability model. *Perception and Psychophysics* 15:149–58. [EC]
Klahr, D. & Wallace, J. G. (1976) *Cognitive development*. Wiley. [THC]
Klapp, S. T., Marshburn, E. A. & Lester, P. T. L. (1983) Short-term memory does not involve the "working memory" of information processing. *Journal of Experimental Psychology: General* 112:240–64. [JPD]
Krech, D., Rosenzweig, M. R. & Bennett, E. L. (1962) Relations between brain chemistry and problem-solving among rats reared in enriched and impoverished environments. *Journal of Comparative and Physiological Psychology* 55:801–7. [PEV]
Lachman, R., Lachman, J. & Butterfield, E. C. (1979) *Cognitive psychology and information processing: An introduction*. Erlbaum. [JPD]
Laosa, L. (1982) School, occupation, culture, and family: The impact of parental schooling on the parent–child relationship. *Journal of Educational Psychology* 74:791–827. [KES]
Laub, J. H. (1983) Urbanism, race, and crime. *Journal of Research in Crime and Delinquency* 2:183–98. [RAG]
Lesser, G. S., Fifer, G. & Clark, D. (1965) Mental abilities of children from different social class and cultural groups. *Monographs of the Society for Research in Child Development* 30, no. 4. [taARJ]
LeVine, R. A. (1975) *Culture, behavior, and personality*. Aldine. [JPR]
Loehlin, J. C., Lindzey, G. & Spuhler, J. N. (1975) *Race differences in intelligence*. W. H. Freeman. [taARJ, JPR]
Luria, A. R. (1976) *Cognitive development: Its cultural and social foundation*. Harvard University Press. [YHP]
Lynn, R. (1978) Ethnic and racial differences in intelligence: International comparisons. In: *Human variation: The biopsychology of age, race, and sex*, ed. R. T. Osborne, C. E. Noble & N. Weyl. Academic Press. [JPR]
(1982) IQ in Japan and the United States shows a growing disparity. *Nature* 297:222–23. [JPR]
MacArthur, R. A. & Elley, W. B. (1963) The reduction of socioeconomic bias in intelligence testing. *British Journal of Educational Psychology* 33:107–19. [RBC]
McCarthy, G. & Donchin, E. (1981) A metric for thought: A comparison of P3 latency and reaction time. *Science* 211:77–80. [EC]
MacGillivray, I., Nylander, P. P. S. & Corney, G. (1975) *Human multiple reproduction*. Saunders. [JPR]
McGue, M., Bouchard, T. J., Jr., Lykken, D. T. & Feuer, D. (1984) Information processing abilities in twins reared apart. *Intelligence* 8:239–58. [rARJ]
McGurk, F. C. J. (1951) Comparison of the performance of Negro and white high school seniors on cultural and noncultural psychological test questions. Catholic University Press [taARJ]
(1953a) On white and Negro test performance and socioeconomic factors. *Journal of Abnormal and Social Psychology* 48:448–50. [taARJ]
(1953b) Socioeconomic status and culturally-weighted test scores of Negro subjects. *Journal of Applied Psychology* 37:276–77. [taARJ]
(1975) Race differences – twenty years later. *Homo* 26:219–39. [taARJ]
McNemar, Q. (1969) *Psychological statistics*. 4th ed. Wiley. [RAG]
Macphail, E. M. (1982) *Brain and intelligence in vertebrates*. Clarendon Press. [EMM]
(1985) Vertebrate intelligence: The null hypothesis. *Philosophical Transactions of the Royal Society of London* B308:37–51. [EMM]
Malina, R. M. (1979) Secular changes in size and maturity: Causes and effects. *Monographs of the Society for Research in Child Development*. Vol. 44, serial no. 179, nos. 3–4. [JPR]
Martin, N. G., Eaves, L. J. & Eysenck, H. J. (1977) Genetical, environmental and personality factors influencing the age of first sexual intercourse in twins. *Journal of Biosocial Science* 9:91–97. [JPR]
Matarazzo, J. D. (1972) *Wechsler's measurement and appraisal of adult intelligence*. 5th ed. Williams & Wilkins. [rARJ, EWPS]
Mercer, J. R. (1984) What is a racially and culturally nondiscriminatory test? In: *Perspectives on bias in mental testing*, ed. C. R. Reynolds & R. T. Brown. Plenum. [RAG, taARJ]
Misawa, G., Motegi, M., Fujita, K. & Hattori, K. (1984) A comparative study of intellectual abilities of Japanese and American children on the Columbia Mental Maturity Scale (CMMS). *Personality and Individual Differences* 5:173–81. [JPR]
Mulaik, S. A. (1972) *The foundations of factor analysis*. McGraw-Hill. [taARJ]
Nagoshi, C. T., Johnson, R. C., DeFries, J. C., Wilson, J. R. & Vandenberg, S. G. (in press) Group differences and first principal component loadings in the Hawaii Family Study of Cognition: A test of the generality of "Spearman's hypothesis." *Personality and Individual Differences*. [RCJ, JRW]
Nettelbeck, T. & Kirby, N. H. (1983) Measures of timed performance and intelligence. *Intelligence* 7:39–52. [taARJ, TN]
Nichols, P. L. (1972) *The effects of heredity and environment on intelligence test performance in 4 and 7 year old white and Negro sibling pairs*. Doctoral dissertation, University of Minnesota. [RAG, taARJ, KES]
Niswander, K. R. & Gordon, M. (1972) *Women and their pregnancies*. W. B. Saunders. [JPR]
Noble, C. E. (1969) Race, reality, and experimental psychology. *Perspectives in Biology and Medicine* 13:10–30. [JGB, rARJ]
(1978) Age, race, and sex in the learning and performance of psychomotor skills. In: *Human variation: The biopsychology of age, race, and sex*, ed. R. J. Osborne, C. E. Noble & N. Weyl. Academic Press. [rARJ]
Occam, William (1320/1965) In: *Gulielmi Ockham opera omnia philosophica et theologica*, ed. E. Moody. Franciscan Institute. [PDB]
Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics) (1982) Profile of American youth: 1980 nationwide administration of the Armed Services Vocational Aptitude Battery. Department of Defense. [RAG, taARJ]
Ogbu, J. (1982) Societal forces as a context of ghetto children's school failure. In: *The language of children reared in poverty*, ed. L. Feagans & D. Farran. Academic Press. [KES]
Olson, R. K., Kliegel, R., Davidson, B. J. & Foltz, G. (1984) Individual and developmental differences in reading ability. In: *Reading research: Advances in theory and practice*, vol. 4, ed. T. G. Waller & E. MacKinnon. Academic Press. [THC]
Omanson, R. C. (1985) Knowing words and understanding texts. In: *New Directions in Child Development #27: The development of reading skills*, ed. T. H. Carr. Jossey-Bass. In press. [THC]
Ombrédane, A., Robaye, F. & Plumail, H. (1956) Résultats d'une application répétée du matrix-couleur à une population de Noirs Congolai [Results of a repeated administration of the coloured matrices to a population of

Congolese Blacks]. *Bulletin du Centre d'Etudes et Recherches Psychotechniques* 5:129–47. [YHP]

Osborne, R. T. & McGurk, F. C. J., eds. (1982) *The testing of Negro intelligence*, vol. 2. Foundation for Human Understanding. [taARJ]

Pachella, R. G. (1974) The interpretation of reaction time in information-processing research. In: *Human information processing: Tutorials in performance and cognition*, ed. B. H. Kantowitz. Erlbaum. [TN]

Palinscar, A. S. & Brown, A. L. (1984) Reciprocal teaching of comprehension-fostering and monitoring activities. *Cognition and Instruction* 1:117–75. [JGB]

Parmellee, A. & Sigman, M. (1983) Perinatal brain development and behavior. In: *Handbook of child psychology, infancy and developmental psychobiology*, vol. 2, ed. M. M. Haith & J. J. Campos. Wiley. [JSC]

Paul, S. (1985) Speed of information processing: The semantic verification test and general mental ability. Doctoral dissertation, University of California, Berkeley. [rARJ]

Payne, R. B. & Turkat, I. D. (1982) Sex, race, and psychomotor reminiscence. *Bulletin of the Psychonomic Society* 19:336–38. [rARJ]

Pearson, P. (1984) *Handbook of reading research*. Longman. [KES]

Pfefferbaum, A., Wenegrat, B. G., Ford, J. M., Roth, W. T. & Kopell, B. S. (1984) Clinical application of the P3 component of event-related potentials. 2. Dementia, depression and schizophrenia. *Electroencephalography and Clinical Neurophysiology* 59:104–24. [EC]

Platt, J. R. (1964) Strong inference. *Science* 146:347–52. [EC]

Plomin, R. (1983) Developmental behavioral genetics. *Child Development* 54:253–59. [KES]

Polich, J., Howard, L. & Starr, A. (1984) P300 latency correlates with digit span. *Psychophysiology* 20:665–69. [EC]

Poortinga, Y. H. (1971) Cross-cultural comparison of maximum performance tests: Some methodological aspects and some experiments with simple auditory and visual stimuli. *Psychologia Africana Monograph Supplement*, no. 6. [YHP]

(1983) Psychometric approaches to intergroup comparison: The problem of equivalence. In: *Human assessment and cultural factors*, ed. S. H. Irvine & J. W. Berry. Plenum. [YHP]

Posner, M. I. (1966) Components of skilled performance. *Science* 152:1712–18. [taARJ]

(1978) *Chronometric explorations of mind*. Erlbaum. [THC, taARJ]

(1982) Cumulative development of attentional theory. *American Psychologist* 37:168–79. [taARJ]

Posner, M. I. & McLeod, P. (1982) Information processing models: In search of elementary operations. *Annual Review of Psychology* 33:447–514. [THC]

Posner, M. I., Pea, R. & Volpe, B. (1982) Cognitive neuroscience: Developments toward a science of synthesis. In: *Perspectives on mental representations*, ed. J. Mehler, E. Walker & M. Garrett. Erlbaum. [MIP]

Pressey, S. L. & Teter, G. F. (1919) A comparison of colored and white children by means of a group scale of intelligence. *Journal of Applied Psychology* 3:277–82. [taARJ]

Price, G. (1984) Mnemonic support and curriculum selection in teaching by mothers: A conjoint effect. *Child Development* 55:659–68. [KES]

Rabbitt, P. M. A. (1979) Current paradigms and models in human information processing. In: *Human stress and cognition: An information processing approach*, ed. V. Hamilton & D. M. Warburton. Wiley. [TN]

(1981) Cognitive psychology needs models for changes in performance with old age. In: *Attention and Performance* 9, ed. J. Long & A. Baddeley. Erlbaum. [MIP]

& Rodgers, M. (1965) Age and choice between responses in a self-paced repetitive task. *Ergonomics* 8:435–44. [PMAR]

Ramey, C. T., Campbell, F. A., & Finkelstein, N. W. (1984) Course and structure of intellectual development in children at high risk for developmental retardation. In: *Learning and cognition in the mentally retarded*, ed. P. Brooks, R. Sperber, & C. McCauley. Erlbaum.

Reuning, H. (1972) Psychological studies of Kalahari Bushmen. In: *Mental tests and cultural adaptation*, ed. L. J. Cronbach & P. J. D. Drenth. Mouton. [taARJ]

Reynolds, C. R. & Gutkin, T. B. (1981) Multivariate comparison of the intellectual performance of blacks and whites matched on four demographic variables. *Personality and Individual Differences* 2:175–80. [RAG, taARJ]

Reynolds, C. R. & Jensen, A. R. (1983) WISC-R subscale patterns of abilities of blacks and whites matched on Full Scale IQ. *Journal of Educational Psychology* 75:207–14. [taARJ]

Rock, D. A., Werts, C. E. & Flaugher, R. L. (1978) The use of analysis of covariance structures for comparing the psychometric properties of multiple variables across populations. *Multivariate Behavioral Research* 13:403–18. [JGB]

Roth, E. (1964) Die geschwindigkeit der Verarbeitung von Information und ihr Zusammenhang mit Intelligenz. *Zeitschrift für experimentelle und angewandte Psychologie* 11:616–22. [TN]

Rushton, J. P. (1984a) Sociobiology: Toward a theory of individual and group differences in personality and social behavior. In: *Annals of theoretical psychology*, vol. 2, ed. J. R. Royce & L. P. Mos. Plenum. [JPR]

(1984b) Group differences, genetic similarity, and the importance of personality traits. In: *Annals of theoretical psychology*, vol. 2, ed. J. R. Royce & L. P. Mos. Plenum. [JPR]

(1984c) Differential K Theory: The sociobiology of individual differences. *Personality and Individual Differences* 6. [JPR]

Ryan, E. (1981) Identifying and remediating failures in reading comprehension: Toward an instructional approach for poor comprehenders. In: *Reading research: Advances in theory and practice*, vol. 3, ed. T. Waller & G. MacKinnon. Academic Press. [KES]

Salthouse, T. A. & Somberg, B. L. (1982) Skilled performance: Effects of adult age and experience or elementary processes. *Journal of Experimental Psychology: General* 111:176–207. [TN]

Sameroff, A. & Seifer, R. (1983) Familial risk and child competence. *Child Development* 54:1254–68. [KES]

Samuels, S. (1979) The method of repeated readings. *Reading Teacher* 32:403–8. [KES]

Sanders, A. F. (1983) Towards a model of stress and human performance. *Acta Psychologica* 53:61–97. [EC]

Sandoval, J. (1982) The WISC-R factoral validity for minority groups and Spearman's hypothesis. *Journal of School Psychology* 20:198–204. [RAG, taARJ]

Sandoval, J. & Millie, M. P. W. (1980) Accuracy of judgments of WISC-R item difficulty for minority groups. *Journal of Consulting and Clinical Psychology* 48:249–53. [taARJ]

Scarr, S. (1981a) Implicit messages: A review of *Bias in mental testing*. *American Journal of Education* 89:330–38. [KES]

(1981b) *Race, social class, and individual differences in IQ*. Erlbaum. [RAG, taARJ, KES]

Scarr, S., Caparulo, B. K., Ferdman, B. M., Tower, R. B. & Caplan, J. (1983) Developmental status and school achievements of minority and non-minority children from birth to 18 years in a British Midlands town. *British Journal of Developmental Psychology* 1:31–48. [JPR]

Scarr, S. & Carter-Saltzman, L. (1982) Genetics and intelligence. In: *A handbook of human intelligence*, ed. R. J. Sternberg. Cambridge University Press. [CB]

Schafer, E. W. P. (1982) Neural adaptability: A biological determinant of behavioral intelligence. *International Journal of Neuroscience* 17:183–91. [taARJ, EWPS]

(1984) Habituation of evoked cortical potentials correlates with intelligence. *Psychophysiology* 21:597. [EWPS]

Schafer, E. W. P. & Marcus, M. M. (1973) Self-stimulation alters human sensory brain responses. *Science* 181:175–77. [EWPS]

Schechter, G. & Callaway, E. (1984) Attention selectively modulates parallel visual processes. Abstract. Society for the Neurosciences, 14th Annual Meeting, vol. 10, part 2. [EC]

Schiff, M., Duyme, M., Dumaret, A. & Tomkiewicz, S. (1982) How much *could* we boost scholastic achievement and IQ scores? A direct answer from a French adoption study. *Cognition* 12:165–96. [KES]

Schmid, J. & Leiman, J. M. (1957) The development of hierarchical factor solutions. *Psychometrika* 22:53–61. [taARJ]

Schönemann, P. H. (1981) Factorial definitions of intelligence: Dubious legacy of dogma in data analysis. In: *Multidimensional data representations: When and why*, ed. I. Borg. Mathesis Press. [PHS]

(1983) Do IQ tests really measure intelligence? *Behavioral and Brain Sciences* 6:311–15. [PHS]

Schwartz, R. M. & Stanovich, K. E. (1981) Flexibility in the use of graphic and contextual information by good and poor readers. *Journal of Reading Behavior* 13:263–69. [THC]

Sen, A., Jensen, A. R., Sen, A. K. & Arora, I. (1983) Correlation between reaction time and intelligence in psychometrically similar groups in America and India. *Applied Research in Mental Retardation* 4:139–52. [taARJ]

Serpell, R. (1979) How specific are perceptual skills? *British Journal of Psychology* 70:365–80. [YHP]

Shucard, D. & Horn, J. (1972) Evoked cortical potentials and measurement of human abilities. *Journal of Comparative and Physiological Psychology* 78:59–68. [taARJ]

Shuey, A. M. (1966) *The testing of Negro intelligence*. 2nd ed. Social Science Press. [taARJ]

Silverstein, A. B. (1980a) Estimating the general factor in the WISC and the WAIS. *Psychological Reports* 46:189–90. [taARJ]

(1980b) Estimating the general factor in the WISC-R. *Psychological Reports* 47:1185–86. [taARJ]

Sisco, F. H. (1982) Sex differences in the performance of deaf children on the WISC-R Performance Scale. Doctoral dissertation, University of Florida. University Microfilms no. DA 8226432. [taARJ]

Sharp, D. M. (1984) Inspection time, decision time and visual masking. Thesis, Department of Psychology, University of Aberdeen. [CB]

Sherry, D. F. (1984) What food-storing birds remember. *Canadian Journal of Psychology* 38:304–21. [EMM]

Sörbom, D. (1974) A general method for studying differences in factor means and factor structures between groups. *British Journal of Mathematical and Statistical Psychology* 27:229–39. [J-EG]

Spearman, C. (1904) General intelligence, objectively determined and measured. *American Journal of Psychology* 15:201–93. [RBC]

(1923) *The nature of "intelligence" and the principles of cognition.* Macmillan. [tarARJ]

(1927) *The abilities of man.* Macmillan. [tarARJ, RCN]

Stankov, L. (1983) The role of competition in human abilities revealed through auditory tests. *Multivariate Behavioral Research.* Monograph no. 83-1. [taARJ]

Stanovich, K. E. & West, R. F. (1979) The effect of orthographic structure on the word search performance of good and poor readers. *Journal of Experimental Child Psychology* 28:258–67. [THC]

Stanovich, K. E., West, R. F. & Feeman, D. (1981) A longitudinal study of sentence context effects in second-grade children. *Journal of Experimental Child Psychology* 32:185–99. [THC]

Sternberg, R. J. (1977) *Intelligence, information processing and analogical reasoning: The componential analysis of human abilities.* Erlbaum. [PK]

(1982) *Handbook of human intelligence.* Cambridge University Press. [HJE, PMAR]

(1984a) *Beyond IQ: A triarchic theory of intelligence.* Cambridge University Press. [JGB]

(1984b) Toward a triarchic theory of human intelligence. *Behavioral and Brain Sciences* 7:269–87. [THC, KES]

Sternberg, R. J., Conway, B. E., Ketron, J. R. & Bernstein, M. (1981) People's conceptions of intelligence. *Journal of Personality and Social Psychology: Attitudes and Social Cognition* 41:37–55. [taARJ]

Sternberg, R. J. & Gardner, M. K. (1982) A componential interpretation of the general factor in human intelligence. In: *A model for intelligence,* ed. H. J. Eysenck. Springer. [tarARJ]

Sternberg, R. J. & Powell, J. S. (1983) Comprehending verbal comprehension. *American Psychologist* 38:878–93. [taARJ]

Sternberg, S. (1966) High speed scanning in human memory. *Science* 153:652–54. [PMAR]

(1969) The discovery of processing stages: Extensions of Donders' method. In: *Attention and performance II,* W. G. Koster. *Acta Psychologica* 30:276–315. [THC, PMAR]

Straumanis, J. J., Jr., Shagass, C. & Overton, D. A. (1973) Auditory evoked response in young adults with Down's syndrome and idiopathic mental retardation. *Biological Psychiatry* 6:75–90. [EC]

Teichner, W. H. & Krebs, M. J. (1974) Laws of visual choice reaction time. *Psychological Review* 81:75–98. [TN]

Thomas, B. (1984) Early toy preferences of four-year-old readers and nonreaders. *Child Development* 55:424–30. [KES]

Thurstone, L. L. (1938) *Primary mental abilities.* University of Chicago Press. [RCN]

Tierney, R. & Cunningham, J. (1984) Research on teaching reading comprehension. In: *Handbook of reading research,* ed. P. Pearson. Longman. [KES]

Trotman, F. (1977) Race, IQ, and the middle class. *Journal of Educational Psychology* 69:266–73. [KES]

Tulkin, S. (1968) Race, class, family, and school achievement. *Journal of Personality and Social Psychology* 9:31–37. [KES]

Undheim, J. O. (1981a) On intelligence. 1. Broad ability factors in 15-year-old children and Cattell's theory of fluid and crystallized intelligence. *Scandinavian Journal of Psychology* 22:171–79. [rARJ]

(1981b) On intelligence. 2. A neo-Spearman model to replace Cattell's theory of fluid and crystallized intelligence. *Scandinavian Journal of Psychology* 22:181–87. [J-EG, rARJ]

(1981c) On intelligence. 4. Toward a restoration of general intelligence. *Scandinavian Journal of Psychology* 22:251–65. [rARJ]

U.S. Department of Labor, Manpower Administration (1970) *Manual for the USES General Aptitude Test Battery.* U.S. Employment Service. [RAG, taARJ]

Van de Vijver, F. J. R. & Poortinga, Y. H. (1982) Cross-cultural generalization and universality. *Journal of Cross-Cultural Psychology* 13:387–408. [YHP]

Vavrus, L. G., Brown, T. L. & Carr, T. H. (1983) Component skill profiles of reading ability: Variations, tradeoffs, and compensations. Paper presented at the meeting of the Psychonomic Society, November 1983, San Diego, Calif. [THC]

Vernon, P. A. (1981a) Level I and Level II: A review. *Educational Psychologist* 16:45–64. [taARJ]

(1981b) Reaction time and intelligence in the mentally retarded. *Intelligence* 5:345–55. [taARJ]

(1983) Speed of information processing and general intelligence. *Intelligence* 7:53–70. [tarARJ, TN, PMAR]

Vernon, P. A. & Jensen, A. R. (1984) Individual and group differences in intelligence and speed of information processing. *Personality and Individual Differences* 5:411–23. [tarARJ, TN]

Vernon, P. E. (1979) *Intelligence, heredity, and environment.* Freeman. [PMAR]

(1982) *The abilities and achievements of Orientals in North America.* Academic Press. [JPR]

Veroff, J., McClelland, L. & Marquis, K. (1971) Measuring intelligence and achievement motivation in surveys. Final report to U.S. Dept. of HEW, OEO, contract no. OEO-4180. Survey Research Center, Institute for Social Research, University of Michigan. [taARJ]

Viaud, G. (1960) *Intelligence: Its evolution and forms.* Hutchinson. [rARJ]

Wade, M., Hoover, J. & Newell, K. (1984) Training reaction and movement times of moderately and severely mentally retarded persons in aiming movements. *American Journal of Mental Deficiency* 89:174–79. [KES]

Wallach, M. & Wallach, L. (1979) Helping disadvantaged children learn to read by teaching them phoneme identification skills. In: *Theory and practice of early reading,* vol. 3, ed. L. Resnick & P. Weaver. [KES]

Wechsler, D. (1958) *The measurement and appraisal of adult intelligence.* 4th ed. Williams & Wilkins. [taARJ]

Weinrich, J. D. (1977) Human sociobiology: Pair bonding and resource predictability (Effects of social class and race). *Behavioral Ecology and Sociobiology* 2:91–118. [JPR]

Weiss, V. (1984) Psychometric intelligence correlates with interindividual different rates of lipid peroxidation. *Biomedica Biochimica Acta* 6:755–63. [EC]

Welford, A. T. (1980) *Reaction times.* Academic Press. [YHP]

Werner, H. & Kaplan, E. (1952) The acquisition of word meanings: A developmental study. *Monographs of the Society for Research in Child Development,* no. 51. [taARJ]

Wherry, R. J. (1959) Hierarchical factor solutions without rotation. *Psychometrika* 24:45–51. [taARJ]

Whimbey, A. (1975) *Intelligence can be taught.* Dutton. [AW]

(1981) *Competency in comprehension (and scientific literacy) as reasoning.* Washington, D.C.: Fund for the Improvement of Post-secondary Education. [AW]

(1983) *Analytical reading and reading.* Innovative Sciences. [AW]

(1984) *Mastering reading through reasoning.* Innovative Sciences. [AW]

Whimbey, A. & Lochhead, J. (1982) *Problem solving and comprehension: A short course in analytical reasoning.* Franklin Institute Press. [AW]

Willerman, L. (1973) Activity level and hyperactivity in twins. *Child Development* 44:288–93. [JPR]

Williams, J. (1980) Teaching decoding with an emphasis on phoneme analysis and phoneme blending. *Journal of Educational Psychology* 72:1–15. [KES]

Wilson, E. O. (1975) *Sociobiology: The new synthesis.* Harvard University Press. [JPR]

Wilson, R. S. (1983) The Louisville Twin Study: Developmental synchronies in behavior. *Child Development* 54:298–316. [JPR]

Wilson, R. & Matheny, A. (1983) Mental development: Family environment and genetic influences. *Intelligence* 7:195–215. [KES]

Wober, M. (1974) Towards an understanding of the Kiganda concept of intelligence. In: *Culture and cognition: Readings in cross-cultural psychology,* ed. J. W. Berry & P. R. Dasen. Methuen. [JSC]

Yerkes, R. M. & Dodson, J. D. (1908) The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology* 18:458–82. [rARJ]