

A Socialization-Mediated Self-Defeating Model for Evaluation of Black-White Performance Differences on Intelligence Test Batteries

JINESH SHAH

Black children are strikingly overrepresented in special educational programs. For example, 43 percent of the black children in New Jersey's public school system were classified as Educable Mentally Retarded (EMR) at a time when black children comprised only 17.8 percent of the public school population.¹ Disproportionately large numbers of black children have thus been placed in special instructional tracks, a phenomenon not exclusive to the New Jersey educational system. Serving as the basis for cognitive ability labels such as EMR, psychometric intelligence tests are the preponderant diagnostic tool for assessing childhood intelligence and consistently yield substantial black-white performance differences favoring whites.² Cognitive measures that reduce this black performance disadvantage have been developed, including the Kaufman Assessment Battery for Children (K-ABC). Although the K-ABC appears to be a more racially equitable battery, it is argued that the K-ABC is not a better psychometric test of mental ability because it does not sufficiently measure a purported general intelligence factor, or g .³

First, I will confirm that classical intelligence tests such as the Wechsler Intelligence Scale for Children-Revised (WISC-R) and Stanford-Binet do in fact assess this so-called g factor more than the K-ABC; that is, they are more highly g loaded. Second, however, I will discuss the nature and calculation of a test's g loading and illustrate that highly g loaded mental measures assess critical relational thinking and abstract cognitive processing abilities. Third, I will argue that black children are likely to fail these types of tasks because their socialization experiences do not include a maternally created "zone of proximal development" that prepares them for the mental manipulations required by highly g loaded tests. I propose that once this failure occurs, anxiety effects interfere with subsequent black test performance.

That is to say, fourth, I will demonstrate that black children enter the testing milieu with high anxiety due to their black-specific socialization experiences. As a result of this high test anxiety, once blacks fail a number of test items that are highly g loaded, they perform even more poorly on subsequent items because content-specific difficulties are compounded by high anxiety-mediated interference. More precisely, fifth, I will show that black children experience a paramount desire to exit the high anxiety testing environment rather than persist trying on subsequent-to-failure test items. Attribution theory will allow us to conclude that the internal attributions that black children generate for their failure may particularly hinder their continued exertion of trying behaviors. Such lack of persistence is of course self-defeating because it will proffer black children with depressed intelligence quotient (IQ) scores, thereby creating the black-white performance disparity. Since the proposed original cause for this black performance disadvantage is the high g

loading of certain tests, I will end by arguing that interpretation of the large black-white gap on highly g loaded mental batteries is dangerous without adequately considering the different socialization patterns of blacks and whites.

To summarize, I propose the following *socialization-mediated self-defeating model* to explain black children's intelligence test performance: encountering highly g loaded items \rightarrow poor performance on these mentally challenging items \rightarrow ascent of anxiety-mediated interference \rightarrow desire to exit from the high anxiety testing milieu \rightarrow low persistence on subsequent items through disengagement from test-taking. Because I define black children's failure on highly g loaded mental tasks as the ultimate origin for their self-defeating low motivation, after attempting to prove the links that my model proposes, I will argue that highly g loaded classical measures such as the WISC-R and Stanford-Binet are not adequate for assessing black children's cognitive capacity. However, more poorly g loaded measures such as the K-ABC may be biased against whites and thus do not offer a suitable alternative to the classical batteries. I will conclude by offering a potential resolution to this dilemma: combining qualitative with quantitative assessment procedures in order to evaluate a child's mental ability with maximal accuracy.

Let me place the implications of my model in perspective. The proposed anxiety-mediated handicapping of black children may be especially dire because elevated test anxiety levels are associated with higher performance for high intelligence students; in contrast, higher anxiety is coupled with a depression of the mental test scores of lower intelligence students.^{4,5} Thus, even if black children are in fact less cognitively capable than their white peers, this black-white mental disparity may be dangerously amplified due to the anxiety-mediated interference that high anxiety blacks may experience during test-taking. In terms of the specific nature of the anxiety components of my self-defeating model, I will show that black children enter the formal test-taking milieu with high test anxiety due to such psychosocial factors as black maternal discipline and evaluation techniques and black stereotype anxiety. I will demonstrate that in order to escape from this anxiety-loaded testing environment—an environment that becomes particularly distressing specifically after the proposed black failure on highly g loaded items—black children are likely to fail to persist after encountering a number of difficult test items. As a result, black children obtain poor test results.

I do *not* contend that white children find no test items challenging. However, I will illustrate that they are more accustomed to the abstract relational thinking needed to succeed on highly g loaded tests and are less "cognitively rigid" because of their socialization experiences. Highly g loaded measures are thus not the extremely debilitating agents that they are for blacks. I will further argue that even when white children do encounter sufficiently difficult test items, they are more motivated to persist trying after their probable failure on these items because

of low test anxiety (unlike black children) and external-to-self causal attributions for their difficulties. As a point of embarkment for analyzing my model, consider a study conducted by Zigler and Butterfield indicating that anxiety-mediated motivational levels do in fact play a central role in governing intelligence test performance.⁶ They demonstrated that increases in children's IQ scores after attending nursery school were due largely to the reduction of crippling anxiety-based negative motivational factors such as a general wariness of adult-run environments like the testing milieu. However, before we begin considering the data that support my model, let us first examine the magnitude of black-white performance differences on various mental tests.

The Black-White Performance Gap is Largest on Classical Intelligence Tests

Converging empirical evidence indicates that black children score approximately one standard deviation (about 15 IQ points) lower than white children on classical standardized tests of intelligence such as the Stanford-Binet and the WISC-R.^{7, 8} In fact, this one standard deviation difference appears on the Stanford-Binet in samples as young as three years of age.^{9, 10} Using Naglieri and Jensen's test data from a sample of fourth- and fifth-graders, we can compare this "classical" black-white gap to the racial performance disparity manifested on the K-ABC.¹¹

The K-ABC's Mental Processing scale (eight constituent subtests) is utilized to calculate an examinee's K-ABC IQ.¹² For raw test scores, the black-white difference (considered in standard deviations) in K-ABC Mental Processing composite scores was only 57% ($[0.42 \sigma/0.73 \sigma] \times 100$) of that exhibited on the Full Scale WISC-R; for standardized scores, only 75% ($[0.58 \sigma/0.77 \sigma] \times 100$).¹¹ (The Full Scale WISC-R score is a summary measure used to compute a WISC-R IQ score that consists of verbal and performance composites.) There was a pronounced contraction of the black-white gap from 9.1 IQ points on the WISC-R to 6.0 IQ points on the K-ABC.¹¹ Is the K-ABC therefore a more racially equitable battery than classical cognitive measures such as the WISC-R and Stanford-Binet? Let us formulate an operational definition of test bias in order to explore this possibility.

The K-ABC Suffers Less from Racial Test Bias than the Classical Mental Batteries

A psychometric test is racially biased when it fails to assess the same abilities in different racial groups; empirical proof of such bias requires demonstrating that the test items (or subtests) do not have the same relative difficulty for the groups.^{9, 13-15} Statistical evidence for racially asymmetric rank orderings of items by difficulty usually takes two forms: (1) a low correlation between the rank ordering of the test items according to their relative difficulty for one group's test data and another group's comparable rank ordering, or (2) a significant racial group \times items analysis of variance (ANOVA) interaction.¹⁶ Applying these bias detection techniques to the data for classical measures such as the WISC-R and Stanford-Binet reveals that these tests suffer from severe racial bias.⁹⁻¹¹ Moreover, given that subtest scores on classical batteries such as the WISC-R correlate more highly for *lower* IQ examinees and blacks are indeed the ones achieving lower IQ scores on these assessment measures, standardly employed cognitive tests will consistently (in a virtually categorical fashion) proffer blacks with poor cognitive competency ratings.¹⁷ In other words, appar-

ently few (if any) of the constituent subtests comprising classical batteries are immune from the racial bias problem due to their high intercorrelation.

How does the K-ABC stand up to this bias litmus test? The K-ABC data do not evidence the same profound corruption by racial bias that afflicts the WISC-R and Stanford-Binet.¹² Is the K-ABC then the assessment tool of choice? A number of psychologists, led by the psychometrician Jensen, would answer with an emphatic "no." Jensen has argued that the K-ABC only yields a reduced black-white difference because it measures a "general" intelligence factor, or *g* factor, to a lesser extent than other psychometric tests; consequently, he concludes that the K-ABC is not a superiorly constructed test, but rather a battery that obtains greater racial parity by including particular subtests that do not sufficiently assess *g*.³ In order to understand this claim, let us consider the nature of the *g* factor.

The General Intelligence Factor *g*: Its Nature and Calculation

Spearman suggested the hypothetical construct of a general intelligence factor *g* which represents the largest common factor assessed by all mental tests.¹⁸ This *g* factor attempts to account for the positive intercorrelations among all cognitive tests. Having observed that all mental tests are in fact positively intercorrelated, Spearman argued that the underlying common ability that they measure is "*g*," a *general* mental capacity due to its ubiquity. "A working definition of intelligence, then, is that it is the *g* factor of an indefinitely large and varied battery of mental tests".¹⁴ In other words, the theoretical construct of *g* is the psychometrician's main measure of cognitive ability. The extent to which a test such as the K-ABC, WISC-R, or Stanford-Binet measures *g*—that is, the correlation between a test and this purported general factor common to all mental measures—is deemed the test's *g* loading.¹⁴

How is a *g* loading actually calculated? The mean of the *g* loadings of a test battery's constituent subtests is the test's overall *g* loading.¹⁴ Recalling the definition of *g* as the "largest common factor," the proportion of the total variance in scores on a subtest shared in common with the battery's other subtests is that subtest's *g* loading. This proportion of shared variance is first determined by factor analysis; the square root of this calculated proportion of common variance then constitutes the subtest's *g* loading.¹⁴ Once the *g* loadings of all the subtests comprising a test battery are calculated, the mean of these loadings is calculated to assign an average *g* loading to the composite battery. Allow me to illumine this *g* loading discussion with a concrete example. Consider a hypothetical matrix of correlations between the scores on four subtests constituting a test battery. Each of the four subtests' *g* loadings can be extracted from this correlational data using factor analysis.

Subtest	A	B	C	D	Row Total
A	[0.63]	0.63	0.51	0.44	2.21
B	0.63	[0.63]	0.49	0.37	2.12
C	0.51	0.49	[0.51]	0.32	1.83
D	0.44	0.37	0.32	[0.44]	1.57
Column Total	2.21	2.12	1.83	1.57	7.73
Factor Loading	0.79	0.76	0.66	0.56	2.78

First, we estimate the so-called “communalities” of the subtests, for these are the (bracketed) values appearing in the matrix’s principal diagonal. A subtest’s communality is the proportion of variance in scores that it shares in common with the other subtests.¹⁴ Since the correlation between two subtests is a measure of their shared variance, to a first approximation a subtest’s communality is the largest correlation that it bears with any other subtest in the correlation matrix.¹⁴ Following this principle, correlation values are entered into the matrix’s principal diagonal.

Second, the row and column values are summed and each subtest’s column total is divided by the square root of the row grand total (7.73)^{1/2}. We divide by the *square root* of the row grand total because the column totals and row grand total are at the r level (both are sums of correlation values), yet a correlation between two subtests measures their proportion of common variance; the *square root* of the correlation between two subtests is the actual correlation of each subtest with the element common to both subtests, or each subtest’s loading on the general factor—the statistic of prime interest.¹⁴ Therefore, we must divide the column totals by the square root of the row grand total to calculate the g loadings that we desire because $r / (r)^{1/2} = (r)^{1/2}$. Doing so, we arrive at our initial crude estimates of each subtest’s loading on the general factor. This first extracted factor, the so-called principal component, is the general factor because it accounts for more of the variance common to the subtests than any other factors that can be subsequently extracted.¹⁴

Third, we extract the second factor common to the subtests. This is accomplished by subtracting from the original values in the correlation matrix that part of each correlation accounted for by the already extracted g factor. Because the correlation between two subtests is the sum of the products of their factor loadings, the correlation between two subtests due to the general factor is the product of their g loadings.¹⁴ For example, the overall correlation between subtests A and B in the original matrix is 0.63. The part of this correlation attributable to the general factor is the product of the g loadings of A and B, or $0.79 \times 0.76 = 0.60$. Consequently, we are left with a residual correlation of $r_{AB} = 0.03$ ($0.63 - 0.60$). We build a residual correlation matrix in this manner. Provided that this matrix is still statistically significant (i.e., r values are not so low that measurement error could account for them), we perform the same operations on this residual matrix that we did on the original matrix to extract the subtests’ loadings on the second factor. Having calculated the loadings on the second factor, we then generate the next residual matrix, extract the loadings on the third factor, and repeat this process until we arrive at a statistically insignificant residual matrix.

Fourth, after extracting all significant factors, we improve our estimates of the subtests’ g loadings. Recall that our first estimate of a subtest’s communality was quite crude (i.e., the largest correlation shared with any other subtest). Since a subtest’s precise communality is the sum of its squared factor loadings, we are now in a position—having calculated the various factor loadings—to recalculate the subtests’ communalities.¹⁴ We incorporate these refined communalities into the principal diagonal of the original correlation matrix and repeat the factor analysis. By reiterating this processing of factor extraction and communality recalculation, we converge on the true communality values of the subtests and thus their true g loadings. Upon calculating each subtest’s true g loading, we *square* this value to determine the amount of each subtest’s variance accounted

for by g ; it then follows that the percentage of the variance in the *composite* test scores accounted for by g is the mean of the squared g loadings of the constituent subtests $\times 100$.¹⁴ We square the g loadings because recall that our g loadings are at the $(r)^{1/2}$ level; however, it is the r form that represents the proportion of shared variance between two subtests, and thus we must transform our g loadings by squaring them. Referring to our hypothetical example and using the crude g loadings obtained after only one round of factor extraction, the percentage of variance accounted for by the general factor g is about 49% ($([0.79^2 + 0.76^2 + 0.66^2 + 0.56^2]/4) \times 100$). Now that we have examined the general factor in depth, let us consider how Jensen uses g loading arguments to attempt to discredit the K-ABC’s utility.

Application of the Spearman Hypothesis to Intelligence Test Data

The concept of g has been used to construct the “Spearman hypothesis” to account for the varying magnitude of the black-white performance gap on different mental tests. The hypothesis contends that a cognitive test’s black-white discriminability is directly proportional to its g loading: measures with higher g loadings yield larger black performance disadvantages.³ In support of this hypothesized relationship, meta-analytic review of the data from 121 cognitive tests indicated that the overall correlation between the test batteries’ g loadings and the respective sizes of their white performance advantage was $r = +0.59$.¹¹

Jensen and his proponents have dismissed the K-ABC’s usefulness by applying the Spearman hypothesis to the K-ABC to conclude that its smaller black-white gap is an artifact of its lower g loading.³ The K-ABC, specifically its Mental Processing scale from which an examinee’s IQ is calculated, *does* indeed have a lower g loading than other standardized intelligence tests. For example, in a population of fourth- and fifth-grade children, the mean g loading of the K-ABC Mental Processing composite was 0.43, lower than the WISC-R subtests’ mean g loading of 0.51.¹¹ In a study conducted by Peoples, Fagan, III, and Drotar, the g loading of the Stanford-Binet was quite high as well, namely 0.68.⁹ Therefore, the K-ABC appears to be less g loaded than classical test batteries. Indeed, specific non- g factors account for 29.0% of the variance on the WISC-R, whereas specific factors account for 41.9% of the variance on the K-ABC.¹¹ Jensen thus concludes that the reduction of the white performance advantage on the K-ABC is a product of the greater specificity of the K-ABC’s test items. Most notably, the K-ABC has a number of subtests which may offset the black-white performance disparity by assessing a non- g specific memory span factor to a large extent.¹¹

Although the g loading data lend support to the Spearman hypothesis, we must carefully consider the implications of a high g loading. Jensen’s denunciation of the K-ABC due to its low g loading places him in the camp that views g as an adequate marker of mental ability. Indeed, Jensen asserts, “We identify intelligence with g . To the extent that a test orders individuals on g , it can be said to be a test of intelligence” (p. 224).¹⁴ Jensen thus favors the use of the Wechsler and Stanford-Binet batteries over the K-ABC because of the higher g loadings of the former tests and poorer g loading of the latter. However, it behooves us to consider the content nature of cognitive measures that are highly g loaded. I intend to argue that a highly g loaded test may be inappropriate for assessing black children’s mental

capacities. Highly *g* loaded tests may foreordain poorer black performance because blacks' socialization experiences do not prepare them for success on these measures (in contrast with white socialization experiences). So allow me to begin exploring this hypothesis with a discussion of the constitution of highly *g* loaded tasks.

Highly *g* Loaded Measures Assess “Noegenetic” and Abstract Mental Ability

Spearman determined that tests highly loaded on the general factor assess cognitive capacities dealing with “noegenesis” and “abstractness” (p. 229).¹⁴ Noegenesis refers to the discerning of relationships among variables and the use of inductive and deductive reasoning. The inventive and creative nature of noegenetic processes is antithetical to the simple imitative and reiterative character of less highly *g* loaded mental tests, tasks that require examinees merely to engage in rote memorization or apply already defined rules (reproductive tasks). Spearman's characterization of highly *g* loaded tests as measures of abstractness alludes to their assessment of an examinee's conceptual thinking power rather than the test-taker's ability to discriminate properties directly through use of sensory perception.

Based on these two defining criteria, highly *g* loaded measures include verbal analogies, sentence completion, and arithmetic problem reasoning; tests with moderate *g* loadings (0.4 to 0.5) include figure recognition and spatial visualization tasks requiring more direct perception and sensorimotor manipulation in order to be completed successfully; finally, rote memorization tasks have extremely low *g* loadings that fall below 0.3.¹⁴ Overall, highly *g* loaded tests require the greatest amount of nonsuperficial cognitive processing or “conscious mental manipulation” (p. 231) by the test-taker.¹⁴ Let me now reiterate my socialization-mediated self-defeating hypothesis in order to place this discussion of highly *g* loaded tasks in the context of my model: black children confront highly *g* loaded test items → probable failure on these items → the rise of anxiety-mediated interference → desire to exit from the distressing testing milieu → low frequency of trying behaviors on subsequent test items and disengagement. I will begin proving my model by using Moore's empirical data on white-socialized and black-socialized adoptees to demonstrate that black-socialized children are in fact less likely to succeed on highly *g* loaded measures due to their lack of training in the critical relational thinking that we have just discovered is crucial for success on these highly *g* loaded tasks.

Black-Socialized Children's Failure on Highly *g* Loaded Measures May be Due to the Lack of a Zone of Proximal Development

Moore examined the socialization patterns of traditionally adopted black children (i.e., black adoptive patterns and thus exposure to black socialization experiences) versus transracially adopted blacks (i.e., white adoptive parents and thus rearing under white socialization experiences).¹⁹ Using this adoption paradigm to partial out that part of the variance in children's mental test scores attributable to socialization variables, a number of studies have attested to the benefit of being raised in a white home environment: the typical one standard deviation poorer performance of blacks on the WISC was virtually eradicated when black children were white-socialized.²⁰⁻²² Moore has expanded the use of

adoption studies to explore specific differences between black and white socialization experiences.¹⁹ The racial differences in maternal behaviors that were identified provide evidence for the hypothesized likely failure of black children on highly *g* loaded items (i.e., those tasks requiring the deduction of complex relationships and abstract, conceptual rationalizing).

The traditional and transracial adoptees completed a “difficult cognitive task” (p. 323), a block design task from the Wechsler Bellevue Scale, with their mothers' assistance.¹⁹ Because this task is rather difficult, it reflects the nature of moderate (if we err on the conservative side) to highly *g* loaded tests and thus allows us to explore black-white socialization-dependent performance differences on measures with appreciable *g* loadings. (For example, the child must determine which shape category each available block belongs to in order to reproduce the block design, and such figure classification tasks have rather high *g* loadings.¹⁴ In addition, a similar design modeling task on the K-ABC, the Triangles subtest, is the most highly *g* loaded K-ABC Mental Processing subtest.¹¹) Moore proceeded to index the mothers' and children's behaviors and verbalizations during problem-solving. Pronounced black-white differences emerged in the domain of maternal helping behaviors.

White mothers provided more nonspecific aid to their children to help them solve the block design problem, whereas black mothers largely provided specific advice.¹⁹ For example, white mothers may have suggested to their children that they work on the problem one section at a time; in contrast, black mothers may have directly instructed their children to orient a block a specific way. Are white mothers then actually less helpful in promoting their children's realization of a difficult cognitive goal such as completion of the block design task? We can apply the developmental theories of Vygotsky to conclude that the white mother's indirect aid may actually be ideal for optimizing a child's success in mental testing situations.

Vygotsky identified a critical factor indispensable to the ontological cognitive development of a child which he termed the “zone of proximal development”.²³ Children interact in a *cognitively nonsuperficial manner* with persons such as their mothers, who function at higher mental levels than themselves, in order to execute mental demands that cannot yet be accomplished independently. This exposure to higher standards of mental functioning is catalytic. In other words, the child strives for and pushes towards acquisition of higher levels of cognitive sophistication as well. The criterion of “cognitively nonsuperficial” participation by the child refers to the fact that the child has to engage deliberately in challenging mental reasoning in order to function cognitively in this mentally complex and advanced milieu (and thereby achieve mental growth). This requisite active mental participation by the child is reminiscent of the high “conscious mental manipulation” required for success on highly *g* loaded tests according to Jensen (see discussion above). This is no accidental coincidence.

White mothers who provide indirect advice to help their children complete difficult mental problems such as the block design task—rather than provide direct, specific aid—may create a zone of proximal development. They compel their children to rationalize problems in a critically analytic manner, to think for themselves and struggle through a cognitive task in order to descry the relationships among the relevant variables and an efficacious problem-solving strategy. Thus,

white mothers may better prepare their children for the mental demands posed by difficult items on cognitive batteries, namely those that are highly *g* loaded. Having been raised in an environment in which they have been mentally challenged by their mothers constantly, white-socialized children are probably prepared to succeed more often on highly *g* loaded test measures.

These white socialization experiences may also foster children's greater independence on mental tasks, whereas black maternal helping behaviors may lead to the development of a pattern of dependency on adults when cognitive demands are particularly rigorous (i.e., tasks not merely assessing simply imitative or rule-applying abilities). In support of this hypothesis, in the Moore study traditionally adopted black children were more likely to request aid from their mothers on the cognitively difficult block design test than the transracial adoptees.¹⁹ Generalizing Moore's findings to home environments—we have no reason to believe that the data lack ecological validity because mother and child were allowed to interact freely during the testing session as they would in the home—black-socialized children have possibly become inured to superficial patterns of thinking. Their mothers have rescued them from deep contemplation and mental wrestling when cognitive demands have become particularly taxing during their socialization experiences. However, there is no mother present to help the black-socialized child on a test battery and therefore he or she is likely to perform poorly on highly *g* loaded tests.

Supporting this zone of proximal development argument, Hess and Shipman have observed a negative correlation of $r = -0.32$ between use of imperative maternal control techniques and children's Stanford-Binet IQ scores.²⁴ Imperative control involves unquestioning obedience of rules and would thus appear to train the child for the rule-applying nature of poorly *g* loaded tasks. In contrast, use of cognitive-rational maternal control techniques is positively correlated ($r = +0.18$) with IQ scores.²⁴ Cognitive-rational control involves providing the child with the logical rationale for rules and inviting critical debate and self-initiated consideration of alternatives by the child, a maternal mode that trains the child in the critical relational thinking about the contingencies between variables that is needed for highly *g* loaded tasks. Although I will concede that these data do not establish causality, they nonetheless strengthen my proposed link between the maternal promotion of complex thinking in children and children's higher achievement on cognitive batteries with high *g* loadings such as the Stanford-Binet.

Moreover, maternal use of cognitive-rational strategies is associated with higher Stanford-Binet behavioral ratings for the child on measures such as activity initiation, response quickness, social confidence, persistence, and eagerness to continue.²⁴ The white maternal socialization mode of creating a zone of proximal development that constantly challenges the child to engage in independent critical thinking—a pattern which we have seen is congruent with the use of cognitive-rational strategies—apparently breeds high self-confidence, high persistence and motivation, and low test anxiety in the child. This finding supports the later steps in my socialization-mediated self-defeating model which posit a connection between failure on highly *g* loaded measures, resulting interference from anxiety effects, and subsequent poor motivation to persist; however, I will explore these links more systematically in subsequent sections.

We can garner further support for the hypothesized poor preparation of blacks for highly *g* loaded tests from Hess and

Shipman's data on differences in maternally supplied aid. Performance scores were lower for children whose mothers provided *direct* physical feedback on a block-sorting task (e.g., showing how to position the blocks) rather than more *indirect* verbal suggestive aid that would facilitate the child's self-discovery of an efficacious sorting principle in a zone of proximal development manner; in addition, there was a negative correlation of $r = -0.35$ between the use of such physical feedback and children's Stanford-Binet IQ scores.²⁴ In contrast, the use of indirect aid—according to my model, a manifestation of a larger maternal socialization pattern of providing a zone of proximal development that forces the child to think critically and abstractly—did not correlate negatively with IQ scores; rather, there was a slight positive correlation.²⁴ Thus, the Hess and Shipman data confirm that low scores on highly *g* loaded tests such as the Stanford-Binet are in fact likely to be obtained by children not benefiting from a zone of proximal development such as blacks. Let me now reiterate the different content nature of highly *g* loaded tests such as the Stanford-Binet and more moderately *g* loaded batteries: the former tests assess the ability to discover relationships among variables and arrive at abstract rules, whereas tests with lower *g* loadings assess more concrete abilities of perception and sensorimotor manipulation.

The lower Stanford-Binet IQ scores of children receiving direct physical feedback may thus indicate that this type of physically oriented maternal support may adequately train children for concrete perceptual and sensorimotor tasks, but not for more abstract cognitive tasks—that is, those that are more highly *g* loaded. In contrast, mothers who provide indirect aid that is more verbal and less physical may promote their children's discovery of abstract rules such as an efficient sorting algorithm in the case of the block-sorting task, thereby preparing them for highly *g* loaded cognitive measures that require such inductive conceptual thinking. To summarize my argument, when problem-solving does not possess sufficient cognitive meaning and salience for a child, the child's performance on highly *g* loaded cognitive tests that are mentally challenging will be severely crippled. This scenario applies to the case of the black-socialized child due to the absence of a maternally created zone of proximal development in his or her socialization.

Recalling the schematic of my model (facing items highly loaded on the general factor → high failure rates on these cognitively complex items by blacks → rise to dominance of anxiety-mediated interference → desire to leave the high anxiety testing environment → low persistence and consequent deteriorated test performance), so far we have explored the first two elements, the *g*-related components of my model. Let me now discuss the hypothesized later steps in this developmental scheme, namely anxiety effects and their mediating role in governing black persistence levels in the testing realm. I will actually discuss these factors in the reverse order in which they appear in my model (i.e., first the low persistence and desire to leave the distressing testing milieu phenomena will be considered and then the anxiety-mediated interference). I choose this "retrograde" explanatory route because I want to lead us directly into a discussion of the crucial middle link between the two main divisions of my model: the progression from high failure rates on cognitively complex highly *g* loaded items to the handicapping of performance by anxiety-mediated interference. Allow me to begin examining the test anxiety component of my theory by considering the last part of my model, the hypothesized low persistence of black examinees.

The Low Test Persistence of Black-Socialized Children

In addition to analyzing maternal and child behaviors on the block design task, Moore administered the WISC to the traditionally and transracially adopted blacks and assessed their behaviors and verbalizations. Traditional adoptees made significantly more refusals to work on test items, so-called “not-work responses,” than transracial adoptees.¹⁹ In other words, rather than face the mental challenges of the WISC, the traditional adoptees frequently disengaged (exited) from the tasks. This finding supports my proposed final step in the development of black self-defeating behaviors, namely that black-socialized children are less likely to persist in the testing realm; apparently, black home environment experiences are more conducive to exit behaviors, whereas white socialization experiences are associated with the exertion of more sustained trying behaviors in mental testing situations. Even more revealing, the black-socialized adoptees’ exit not-work responses were predominantly of the negation (direct refusal) and substitution (providing an impertinent utterance rather than engagement in task-specific behavior) form; in contrast, the preponderance of the white-socialized adoptees’ not-work responses were refusals expressed in terms of task-specific competence (e.g., inadequate ability to complete a given test item).¹⁹ We can use attribution theory to understand the implications of these differences between the transracial and traditional adoptees’ exit not-work responses.

The white-socialized adoptees were more apt to attribute their difficulties on test items to a situational external locus such as lack of experience (e.g., “I haven’t learned that yet”) rather than a more invariant dispositional internal locus (e.g., generally poor ability).¹⁹ Such an attribution is supremely adaptive to test-taking because the child will be motivated to endeavor on subsequent test items after failure on prior items because of the intact confidence in his or her capacity to complete successfully items that are suited for his or her base of knowledge. In other words, it is only illogical for one to continue trying when there is little chance of being rewarded with success. Because the white-socialized child has not attributed his or her failure to a stable and unchanging internal locus and therefore retains the potential for success on proceeding items, it is not dissonant for him or her to continue trying on subsequent items in the test battery.

On the other hand, traditional adoptees most likely developed doubts about their abilities because they failed to frame their not-work responses in terms of external-to-self experiential difficulties. As a result, in contrast to the transracial adoptees, the black-socialized children probably became dismayed by their failures. They had low motivation to persist and sought quick test exit (in accordance with the final steps of my socialization-mediated self-defeating model) due to the perceived inevitability of continued failure. In support of this hypothesis, typical not-work responses of the traditional adoptees were shaking the head “no” in response to test items before the tester even completed reading questions and declaring, “I don’t like to do this” (negation responses), or singing a song and playing finger games (substitution responses), all of which were clear attempts to disengage from test-taking rather than persist.¹⁹ These behaviors of the traditional adoptees conform perfectly with those of persons with self-defeating profiles according to Curtis: “they may have high test anxiety and avoid tasks on which they expect to fail...they may attend to distracting features of the situation [substitution behavior] when

they fail to meet defined performance criteria, especially if they become very anxious” (p. 68).²⁵ Curtis is establishing a link between high anxiety and low persistence, and I will explore this relationship below when I consider steps earlier than the terminal “failure to persist” one in my model.

Overall, the implications of the not-work findings are enormous, for typically not-work responses on intelligence tests are treated like incorrect responses. For example, when a subject fails to reply to a number of consecutive items on a Wechsler subtest, the subtest is terminated and the subject is assigned a score of zero for all subsequent unattempted items.²⁶ Moore’s work alerts us to the danger of assuming that not-work responses are indicators of deficiencies in cognitive skills, for the frequency of not-work responses may be a measure of a child’s motivation to persist on intelligence batteries rather than an indicator of deficits in the child’s cognitive repertoire. Allow me now to address my model’s hypothesized black high test anxiety profile and the interference effects that this anxiety wields on black test performance.

Anxiety-Mediated Interference with the Test Performance of High Anxiety Blacks

Supporting the anxiety component of my model, Payne, Smith, and Payne demonstrated that black students obtain higher test anxiety scores than their white peers on the Survey of Feelings About Tests self-report which largely reflects the general factor (as determined by factor analysis) common to the most widely used test anxiety scales such as the Children’s Manifest Anxiety Scale.²⁷ Given this initial foothold for validating the anxiety elements of my model, let us consider the relevant data on the high test anxiety of black students more thoroughly.

Empirical evidence for socialization-dependent differential levels of children’s engagement in the testing environment is quite revealing. Ingeniously, the levels of “spontaneous extension,” or “unsolicited elaborations related to the task at hand” (p. 319), of transracially and traditionally adopted blacks were indexed during performance of the WISC.¹⁹ The frequency of a child’s use of such self-initiated utterances indicates his or her degree of involvement, motivational level, and comfort with the testing environment—in other words, how low (or high in the case of infrequent spontaneous extension behavior) test anxiety is.¹⁹ Therefore, if the socialization-mediated self-defeating hypothesis is correct in asserting that the test performance of black-socialized children is crippled by high anxiety-induced low motivation to persist, then black socialization experiences should proffer lower levels of spontaneous elaboration during test-taking.

Indeed, traditionally adopted blacks provided fewer spontaneous elaborations than their transracial counterparts.¹⁹ White socialization experiences may thus foster the greater ease in testing situations that promotes high engrossment and motivation to excel rather than the desire to exit the test situation as quickly as possible due to high anxiety. The apparent high anxiety of black-socialized children in the testing environment, in contrast, may make them more likely to ascribe failure to a stable internal locus and thus cease trying on subsequent test items: “individuals high in anxiety may be more likely to attribute failure to lack of ability than persons low in anxiety” (p. 76).²⁵ (This statement is consonant with my discussion in the previous section of the detrimental attributions that blacks make for their testing failures.)

Let me emphasize here that my model claims that poor black motivation is *caused* by anxiety-based interference. In fact, empirical evidence demonstrates that this postulated causal link between anxiety-mediated inhibition of the performance of high anxiety children and the compromise of trying behaviors is stronger in the proposed forward direction than in the reverse direction.²⁸

Convergent data corroborate that a prominent causal agent for black children's poorer performance on intelligence tests relative to whites is high test anxiety (which results in the use of self-defeating exit behaviors to escape from the distressing testing milieu according to my model). First, the use of pre-testing anxiety reduction mechanisms reduces the black-white difference by as much as 10 IQ points (a standard deviation is about 15 IQ points).¹⁹ White children do not benefit nearly as much as blacks by such optimization maneuvers, thus indicating that a black-specific deficiency is corrected by such procedures rather than a race-independent general handicapping. One such optimization protocol achieved greater performance parity by using a pre-testing play period to reduce the situational test anxiety of otherwise disadvantaged children.²⁹ Second, Seitz, Abelson, et al. demonstrated that the location of the testing environment can substantially contribute to the debilitation of disadvantaged children's test performance and that high anxiety-based negative motivational factors possessed by these children can be diminished with an anxiety-reduction regimen of one-to-one interaction with adults.³⁰ Third, providing ameliorative intervention to high anxiety children (which would include blacks based on the cited data), such as having them experience success and/or participate in neutral (nonevaluative) tests, substantially improves their performance.³¹ In totality, these studies indicate that the differential anxiety levels of blacks and whites play a central role in governing their test performance. However, we have not yet identified specific "culprits" of the black socialization experience that preclude optimal black achievement in testing situations due to anxiety-mediated performance interference. I intend to demonstrate that the differing nature of white versus black maternal behaviors vis-à-vis their children may account for the corruption of black children's test performance by anxiety effects.

Maternally Determined (and Related) Socialization Experiences Account for the Anxiety-Mediated Interference with Black Children's Intelligence Test Performance

A clear pattern of differences emerged between black and white mothers when they were asked to assist their children with the block design task in the Moore study. Whereas white (transracial) mothers used more positive ways of releasing tension such as joking or laughing when their children espoused unfruitful strategies, black (traditional) mothers used more negative tension-release behaviors such as grinning and scowling.¹⁹ White mothers also coaxed their children to success mainly with positive evaluations of praise and enthusiasm (e.g., urges and cheers); in contrast, black mothers predominantly engaged in negative evaluations of their children's abortive problem-solving efforts and tried to boost performance by expressing their dissatisfaction with present execution patterns.¹⁹ As a result of their preponderant use of positive forms of tension release, positive evaluations of their children's work, and urging with enthusiasm, white mothers may breed in their children higher

levels of confidence in their mental problem-solving capacities and lower anxiety levels during mental task completion than black mothers. Due to "the positive affect and encouragement their mothers display" (p. 325), white-socialized children may have low anxiety and high motivation in problem-solving situations and may thus be likely to persist wholeheartedly (rather than adopt exit behaviors) in their attempts to solve the difficult mental tasks that appear on highly loaded test batteries.¹⁹ Supporting this claim, Crockenberg demonstrated that greater maternal emotional responsiveness is in fact likely to elevate children's motivational levels on cognitive tests.³²

Furthermore, unlike black-socialized children, white-socialized children are likely to have received the low-pressure message, "It's okay to be wrong as long as you are trying" (p. 325); as a result, after failure they may be less likely to engage in self-defeating exit behaviors due to the absence of intense pressure.¹⁹ On the other hand, consider our discussion of black mothers' lack of encouragement of exploratory behaviors that are not necessarily conducive to immediate solving of a problem and their use of negative reinforcement mechanisms. This black maternal socialization pattern may account for the diminished efforts of black-socialized children in response to failure on a test battery (recall the greater frequency of not-work responses among traditionally adopted blacks and their expression of not-work responses largely as negations and substitutions). Indeed, rather than risk the disapproval of their intolerant mothers (or other adults when we generalize to the formal testing milieu) with incorrect test responses, black children would be expected to adopt the strategy of exiting difficult mental tasks (use of negation or substitution). Blau provides corroborating evidence for the effects of racial differences in maternal behaviors on children's test performance: black mothers use much more aversive discipline than their white counterparts, a behavioral difference which may obviate black children's need to reciprocate for positive benefits (which are lacking in the case of black maternal use of aversive discipline) received from the mother-child interaction.³³

Gay and Abrahams confirm the proposed contribution of black maternal behaviors and familial social structure to the anxiety-mediated interference with black children's intelligence test performance.³⁴ The typical black home is matriarchal and the mother largely defines her self-worth by the efficiency of her household's operation.³⁴ In addition, children in the black home are almost exclusively reared by their siblings after they are toddlers.³⁴ This developmental mode causes the black child's rare interactions with adults to occur when he or she disrupts the smooth functioning of the household and the mother must directly query the child to assess why things have gone awry.³⁴ The child may thereby associate direct questioning with the anxiety-invoking allegations of wrongdoing. When this contingency is generalized to the asking of direct questions in the testing situation, children from black households will suffer from high tension in the testing realm, making the desire to exit and the manifestation of self-defeating behaviors highly probable—that is, completion of the sequence of events in my self-defeating model. One may argue that the tense nature of the black child-mother interaction will not necessarily be generalized to the testing milieu. However, the black child's model for interaction with adults is almost exclusively the child-mother relationship due to the virtual absence of any other child-adult interaction in black socialization

experiences. It is thus highly probable that the tense child-mother interaction will be generalized to the child-adult examiner testing interaction.

How may this anxiety-loaded black child-mother interaction lead to the black child's poor test performance? The final "product" of low test persistence in my model can occur when test anxiety is high *and* expectations of success are high.²⁵ We have already discussed how black children do in fact have high test anxiety. Moreover, since the black child may associate the direct questioning characteristic of test-taking with the scenario in which he or she fails to meet the mother's *high expectations* for him or her to contribute to the smooth operation of the household, the latter criterion of high expectations is satisfied as well.

Finally, other features of the black psychosocial environment may aggravate the self-defeating problem by increasing the anxiety level that black children enter the testing situation with. Steele demonstrated that the socialization of black children makes them greatly cognizant of the racial stereotype of their lower intellectual ability; this racial stereotype burden, in turn, predisposes blacks to engage in self-fulfilling prophecies to validate this stereotype.³⁵ Stereotype vulnerability was reduced and greater racial parity achieved in the Steele study only when this "stereotype anxiety" was made less salient.³⁵ An intriguing theory from the literature on self-defeating behaviors illustrates the gravity of this stereotype anxiety handicapping.

Persistence on a test may be inversely proportional to the diagnostic value of the test items.²⁵ In other words, if test items are conceived as opportunities to reduce one's uncertainty about one's cognitive abilities, then on a test with items that greatly reduce this uncertainty (items with high diagnostic value), subjects will stop exerting trying efforts more quickly than on a battery with a lower rate of uncertainty reduction. Allow me to frame black stereotype vulnerability into this model. If the black child enters the testing situation with high stereotype anxiety having used racial stereotypes to conclude already that he or she is almost certainly cognitively inferior to whites, then test items will assume high diagnosticity. That is, the black child will need little convincing about his or her "inferior" mental ability because he or she has made a strong *a priori* evaluation. As a result, when difficult test items such as highly *g* loaded measures on which failure is likely (recall the zone of proximal development argument) are encountered, poor performance on these items may serve as powerful confirmation of the premature conclusion of lesser ability. Black testing persistence will thus be compromised. In contrast, the white child who likely does not enter the testing milieu with a strong predetermined evaluation of his or her intellectual prowess will not artificially inflate the diagnostic value of the test items and will thus persist longer due to the higher informative value of the items.

Let me summarize my arguments thus far. Recall once again my socialization-mediated self-defeating model for black test performance: items with high *g* loadings encountered → shoddy performance on these cognitively complex items → ascent of anxiety-mediated interference → desire to escape from the high tension testing environment → disengagement through low persistence on ensuing items and resulting severe deterioration in test performance. I garnered evidence for the first main part of the model, black children's poor success on highly *g* loaded tests, by demonstrating that they lack a maternally provided zone of proximal development that is crucial for success on the

abstract relational tasks appearing on highly *g* loaded tests. I then provided data supporting the second chief part of my model, the debilitating influence of black children's high test anxiety and their resulting use of self-defeating exit behaviors: black mother's greater use of negative tension and negative evaluations of their children's problem-solving, the black child's association of test-taking with the high anxiety situation in which the mother queries why the household's smooth functioning has been upset, the black child's vulnerability to stereotype anxiety, and other typical black psychosocial experiences as evidence for a high test anxiety profile among black examinees; the fewer spontaneous elaborations (a measure of low test anxiety) of black-socialized children, their more frequent use of exit not-work behaviors, and their framing of not-work responses in terms of internal rather than experiential attributions as evidence for the compromise of black test persistence due to anxiety effects. We are now ready to join these two principal components into the unified model by discovering the crucial link between the two: how failure on mentally challenging highly *g* loaded test items connects with the anxiety-mediated corruption of the high test anxiety black child's subsequent-to-failure test performance.

The Connection between Black Children's Failure on Highly *g* Loaded Measures and Anxiety-Mediated Interference

First of all, it has been observed that children with high anxiety (such as blacks) do in fact experience more failure on mental tests than their low anxiety cohorts.³¹ The data of Leppin, Schwarzer, et al. provide a crucial link between test failure and high test anxiety effects, one that we can apply to blacks due to their demonstrated high anxiety profiles.³⁶ An ingenious paradigm was used to examine *specifically* the effects of test failure on high anxiety examinees: subjects completed anagrams and received categorical failure feedback regardless of actual performance. It was discovered that high test anxiety subjects attributed their difficulties more frequently to inability (a stable internal locus) and less to the difficulty of the cognitive tasks.³⁶ Such a global attribution hinders further trying behaviors because failure is viewed as inevitable. In contrast, low anxiety examinees implicated task difficulty rather than their own lack of ability in their failure.³⁶ Hedl confirmed that high anxiety individuals do in fact believe that their failure is due to global deficits rather than task-specific demands.³⁷

This attribution pattern conforms with Moore's finding that black-socialized (i.e., high test anxiety) children's not-work responses were *not* phrased in terms of external-to-self experiential impossibilities, unlike the white-socialized adoptees' not-work responses. In turn, the traditional adoptees demonstrated lower persistence through their more frequent espousal of exit behaviors. The studies of Leppin, Schwarzer, et al. and Hedl thus corroborate that the high anxiety profile of black-socialized children is in fact associated with anxiety-mediated interference with subsequent-to-failure test performance. In other words, high anxiety blacks will engage in unfavorable causal attributions for their failure that are likely to engender poor continued persistence and thereby deteriorate subsequent performance because of the belief that their low global ability predetermines failure. This corruption of black performance by anxiety elements is confirmed by the fact that high test anxiety children such as blacks have a low "worry"

threshold; that is, they “react with worry at the first hint of failure” (p. 64) according to Hedl.³⁷ Hedl thus verifies that anxiety effects interfere with the performance of high anxiety black children once failure occurs, failure which is likely to occur on highly *g* loaded tests due to the black child’s discussed lack of practice with the abstract cognitive processing necessary for success on these tests. Before considering more evidence, let me emphasize that my model asserts that the detrimental effects of anxiety-mediated interference occur *in response to failure*. In support of this claim, high anxiety children do in fact perform as quickly and accurately as low anxiety children when provided with experiences of success; differential performance has only been observed in the case of failure.³¹

This link between failure and anxiety-mediated interference effects among black children can be viewed not only in terms of blacks’ attributions for failure, but also in terms of the black-white maternal differences in encouragement of exploratory behaviors during problem-solving that we discussed. Recall that unlike white mothers, black mothers discourage their children from exploring different strategies for tackling a mental task if they are not directly conducive to solving the given problem. Black-socialized children may thus grow fond of using traditional, well-tried strategies that have proffered success in their past experiences with mental demands. Such “cognitive rigidity,” however, increases the amount of anxiety-mediated interference with test performance that black-socialized children experience.²⁸

Rigid persons will have difficulty completing many of the rather novel and atypical tasks appearing on test batteries, most notably items which are highly *g* loaded and thus require item-specific creative, abstract relational thinking rather than application of a general cognitive rule or use of imitative, reproductive strategies. Jerusalem, Liepmann, et al. indicate that mentally rigid persons—such as black children—are in fact likely to experience failure more frequently in the testing environment, experience a resulting decrease in self-esteem and confidence in their cognitive abilities, and become inhibited in subsequent performance by anxiety effects.²⁸ Thus, the cognitive rigidity of blacks compounds the anxiety-mediated interference elements that we have already considered. Rigid individuals will also attempt to avoid situations in which their cognitive capacities come under question.²⁸ This claim concurs with my socialization-mediated self-defeating model’s assertion that black-socialized children are more likely to engage in exit behaviors in order to escape from the anxiety-loaded testing environment.

Finally, we can consider the use of aversive discipline by black and white mothers to understand the connection for black-socialized children between test failure and anxiety-mediated interference. Parents of high anxiety children are more likely to use aversive discipline and be critical rather than constructive during their children’s mental task efforts.³¹ High anxiety children may thus develop high dependency on their mothers to deal adequately with many cognitive tasks due to their failure to acquire successful strategies for mental problem-solving. Hence, high anxiety black children are expected to experience high failure in the formal intelligence testing milieu in which the mother is no longer present to help, especially on highly *g* loaded items that test the limits of problem-solving capacity. In contrast, the parents of low anxiety children, such as those of white children, help their children develop “effective problem-solving strategies *without completely taking over the problem solving*...they teach their children to *rely on their own resources* and may help

them learn task-oriented responses, thereby increasing their effectiveness as problem-solvers” (p. 60; italics added).³¹ In other words, white mothers provide their children with the zone of proximal development that hones their mental capacities and prepares them for highly *g* loaded items appearing on mental test batteries. Overall, we have pondered data about black children’s attributions for failure, blacks’ cognitive rigidity, and black maternal use of aversive discipline and the resulting black child’s cognitive dependency to link the failure and anxiety components of my model. We have thus now integrated the two main elements of my model into a unified theory. However, before considering my model’s implications, allow me to delineate the difference between black performance on tests with high versus low *g* loadings.

High Anxiety Blacks are Handicapped Specifically on Highly *g* Loaded Tests

An important distinction that my socialization-mediated self-defeating model makes is between black performance on highly *g* loaded items and tasks with poorer *g* loadings. I contend that high anxiety black children experience interference due to their high anxiety profiles on the former type of items, not on the latter. The literature in fact indicates that “performance on easy (and therefore success-likely) test items shows little or no dependency on anxiety level” (p. 128), but high anxiety subjects do in fact perform poorly relative to low anxiety subjects on more difficult measures.³⁸ Empirical evidence for this distinction between difficult, cognitively complex (highly *g* loaded) test items and relatively easier (less strongly *g* loaded) items can be elicited from the work of Rocklin and Thompson.³⁹

These psychologists constructed two tests consisting of verbal ability items originally appearing on the Standardized Achievement Tests (SATs), a “hard” test consisting of those items with low passing rates on the SAT and an “easy” version comprised of items which more examinees passed. Higher anxiety students performed better on the easy test than on the hard test, and this performance profile was *not* merely an artifact (i.e., performance is expected to deteriorate with increasing item difficulty), for the low anxiety students actually performed better on the hard test!³⁹ As my model proposes, apparently the high anxiety black child is indeed likely to perform poorly relative to whites specifically on difficult highly *g* loaded test items. In fact, when the preferences of high and low test anxiety subjects for easy versus difficult test items were assessed by Arkin and Haugtvedt, there was a marked preference among high anxiety subjects for easier test items.⁴⁰ “Difficult tasks tend to generate prodigious amounts of cognitive interference among test-anxious persons” (p. 159).⁴⁰ Therefore, high test anxiety examinees’ choice of easy items may be indicative of their desire to avoid such unpleasant “prodigious amount of cognitive interference.” This theory is completely consonant with my socialization-mediated self-defeating model: I likewise hypothesize that black children’s failure on highly *g* loaded tasks gives rise to anxiety-mediated interference, and they then attempt to avoid the resulting highly tense and uncomfortable testing situation by using disengaging exit behaviors.

Covington verifies these differing interactions between high versus low task difficulty and interference from anxiety effects.⁴¹ Subjects were asked to perform “convergent” deductive reasoning tasks which “stress the individual’s capacity to generate and then selectively eliminate possible

ideas until the person *converges* on only one idea that fits all the constraints of the task" (p. 102; italics added).⁴¹ Covington's tasks would presumably be highly *g* loaded, for we have discovered that measures which assess rule abstraction capacities and deductive (convergent) reasoning power are in fact highly *g* loaded. Interestingly, Covington found an inverse relationship between test anxiety and performance, one that increased in magnitude with increasing task difficulty.⁴¹ Therefore, test performance interference due to anxiety effects does seem to manifest itself predominantly on more difficult test items, which include highly *g* loaded tasks due to their discussed cognitive complexity. Let us now consider Covington's explanation for this interaction between high anxiety and high item difficulty.

Covington contends that high anxiety examinees fear that they are constantly falling behind on difficult test items; however, this "reverberating circuit of worry" (p. 104) does *not* manifest itself on simpler tasks requiring more superficial cognitive processing such as simple rule-applying or imitative, reproductive behavior: "intrusive worry may create an inhibition of all *but* the simplest, automatic ideational responses, thus causing an inability to focus on completion of a task" (p. 104; italics added).⁴¹ Difficult, cognitively complex highly *g* loaded measures are thus *specifically* the culprits for the self-defeating behavior ("an inability to focus on completion of a task") of high test anxiety black children; anxiety-mediated interference is not expected on more poorly *g* loaded measures that are less demanding and require comparatively perfunctory mental calculations. Furthermore, Covington indicates that the *worry* that one is falling behind during a test—that is, the perception of failure to perform adequately—results in anxiety-mediated "inhibition" of task completion. This theory gives strong support to my model's link between black children's probable failure on highly *g* loaded items and the consequent disruption of subsequent test performance by anxiety effects (i.e., "worry" in Covington's phraseology).

One may argue that my model fails to consider that white children will also eventually encounter on a test battery items that are sufficiently difficult such that they (just like blacks) will experience failure that corrupts subsequent performance via anxiety effects. However, there is strong evidence that this anxiety-mediated interference is a black-specific phenomenon. Consider that failure attribution to an internal locus such as lack of ability is likely to engender anxiety and feelings of helplessness.⁴² Indeed, we have seen that black-socialized children failed to frame their difficulties in the Moore study in terms of experiential impossibilities and instead defined internal loci such as poor global cognitive ability. We discovered that this attribution profile of the black-socialized child is in turn conducive to the development of anxiety-based interference.

We have also observed that white-socialized children are likely to ascribe their test failure to less invariant, more external-to-self causes such as item difficulty or lack of task-specific competence. Test-takers making such unstable attributions are likely to respond to failure as a cognitive challenge and are motivated to try harder because they believe that they can succeed; that is, they experience "facilitating anxiety."⁴² In contrast, black-socialized children who choose more stable internal sources for their test difficulties are likely to perceive failure as threatening rather than challenging; that is, they experience "debilitating anxiety."⁴² This "debilitating anxiety" leads to a loss of the sense of control: the

black-socialized child believes that his or her performance will be inevitably poor due to his or her low ability (i.e., a loss of perceived contingency between effort and outcome); as a result, trying behaviors are compromised. On the other hand, since white-socialized children do not experience this "debilitating anxiety" as a result of test failure, my self-defeating model applying to black-socialized children fails to apply to white-socialized children's performance: test failure does not result in detrimental interference from anxiety effects and thus the chain leading to deteriorated test performance is halted. Now that we have examined my model in full, let us consider the concerns that other psychologists might have regarding my socialization-mediated self-defeating theory.

Addressing the Arguments of Jensen and Herrnstein and Murray

What does Jensen have to say about the issue of test anxiety and its contribution to the black-white gap in IQ scores? Collectively referring to variables such as potentially higher self-confidence and lower test anxiety levels among white children as test sophistication or "test-wiseness" (p. 589), Jensen contends that there is no evidence that test sophistication is a significant contributor to the black-white performance difference on mental tests.¹⁴ He supports this conclusion by citing studies that examine possible black-white differences in the benefits incurred from test practice. (Test practice is a means of increasing test familiarity and thus reducing test anxiety and increasing "test-wiseness.") However, Jensen does not adequately consider the relevant data to the test sophistication issue: the differential availability to blacks and whites of informal practice/training episodes in the non-testing realm.

Jensen focuses on data indicating that if black and white children are given a formal cognitive test a first and then second time, the magnitude of performance improvement is about the same for both groups.¹⁴ Such findings merely indicate that practice elevates performance levels approximately the same for blacks and whites, but fails to address the frequency issue. In other words, Jensen ignores the fact that white socialization environments simulate testing conditions more frequently than black ones because of the discussed central role that the zone of proximal development (in which the mother constantly challenges the child cognitively by not providing direct aid) plays in the white home environment. Since white children are therefore likely to have greater practice functioning under mentally challenging test-like conditions, they will probably have lower test anxiety and thus outperform blacks on mental tests. Essentially, Jensen contemplates relative performance improvements rather than the more revealing absolute performance levels of blacks and whites: although blacks and whites benefit about the same from taking a formal test twice, blacks still achieve appreciably lower absolute scores than their white counterparts due to their pre-testing differential socialization experiences.

An influential work in recent years on the subject of intelligence testing is *The Bell Curve* by Herrnstein and Murray. Contrary to my socialization-mediated self-defeating model, Herrnstein and Murray deny that motivational factors play any substantial role in depressing black children's IQ scores.¹⁶ They support their claim mainly with two bodies of data. First, they examine the performance of blacks and whites on the digit span subtest of the Wechsler, consisting

of a forward digit span test in which the subject has to repeat a series of numbers in the same order than he or she hears them and a backward digit span test in which the number sequence must be recalled in reverse. Although both tests have the same content, the backward digit span test is much more highly *g* loaded and the white advantage is about twice as great on this test as on the forward version.¹⁶ Herrnstein and Murray argue that motivation cannot account for this difference in performance on the two parts of the digit span subtest because both content-identical parts occur at the same time during the Wechsler battery.¹⁶ However, they do not consider the data that I invoked earlier about the probable absence of black self-defeating behaviors on subtests with poor *g* loadings that require rather superficial cognitive processing. The forward digit span is such an exceedingly easy task—indeed, it has half the *g* loading of the backward digit span test³—that anxiety-mediated interference probably does not occur to any appreciable extent on it. In the Covington framework that we have discussed, this forward digit span test is likely a task requiring the “simplest, automatic ideational responses” and thus “inhibition” by “intrusive worry” does not occur on it (p. 104).⁴¹ Only on the significantly more mentally challenging (i.e., more highly *g* loaded) backward digit span test may the high test anxiety of blacks become a liability; hence, the significantly more pronounced black-white gap on the backward digit span test.

The second piece of evidence that Herrnstein and Murray base their conclusions on is work conducted by Jensen. Subjects depressed a “home” button and then moved a finger as quickly as possible to another button above which a light had gone off. Jensen measured both reaction time (the time required to lift the finger off the home button when the target light illuminated) and movement time (the time needed to move the finger from just above the home button to above the target button).¹⁶ Jensen discovered that blacks consistently had slower reaction times but faster movement times, and thus Herrnstein and Murray reason that motivational factors are not important in governing black test performance: “how can one be unmotivated to do well during one split-second of a test but apparently motivated during the next split-second?” (p. 228).¹⁶ However, these results can be rationalized by my self-defeating hypothesis. The quick movement time may *not* be an indication of high motivation *to excel*. Rather, the short movement times of blacks can be viewed as attempts to exit the highly distressing testing situation quickly. After possibly becoming dismayed by the difficulty of the actual cognitive processing component of the task (i.e., deciding which button to move one’s finger to, a mental calculation reflected in the reaction time), the black child was motivated to end his or her involvement in this frustrating, anxiety-inducing situation. In other words, the black child’s quick movement time may be an indication of high motivation *to exit* rather than *excel* in the testing environment.

Overall, therefore, I conclude that both the team of Herrnstein and Murray and the psychometrician Jensen erroneously reject the claim that test anxiety and motivational factors assume a central role in the poorer performance of black children on cognitive test batteries. Having attempted to bolster the validity of my model through this discussion of Jensen’s and Herrnstein and Murray’s probable criticisms, allow me now to suggest the implications of my theory on the use of classical measures such as the WISC-R and Stanford-Binet to assess blacks.

Both Highly and Poorly *g* Loaded Intelligence Tests May be Racially Biased

I have argued that the self-defeating behaviors of black children in the testing environment stem from highly *g* loaded items. We have also observed that classical psychometric batteries such as the WISC-R and Stanford-Binet are more highly *g* loaded than the K-ABC. Should we then discard the WISC-R and Stanford-Binet and adopt the K-ABC for measuring intelligence? Is this an efficacious strategy for preventing black children from engaging in self-defeating behaviors that lead to potential gross underestimation of their mental abilities? The solution is not this easy.

Jensen noted that the K-ABC’s poorer *g* loading relative to classical measures allows it to achieve greater racial parity (see discussion above). Specifically, some of the K-ABC’s subtests (e.g., Numerical Recall) assess the non-*g* specific factor of sequential short-term memory to an appreciable degree. The Naglieri and Jensen data indicate that blacks are not handicapped relative to whites on these more specific and thus more poorly *g* loaded tasks.¹¹ We have discussed how tasks that are not highly *g* loaded (such as rote memorization tasks like the Numerical Recall subtest) are not nearly as cognitively demanding as measures that are in fact highly *g* loaded. On these types of simpler mental tasks, the anxiety factor may most accurately be viewed as “a general increase in drive level which *enhances* the individual’s readiness to respond where only one dominant response tendency exists” (p. 128; italics added).³⁸

Therefore, on the K-ABC’s sequential short-term memory demands that most test-takers may be able to respond to correctly regardless of race (i.e., “only one dominant response tendency exists”), anxiety may be facilitating rather than debilitating. As a result, black children—who have higher anxiety levels than their white peers—do fairly well on the K-ABC subtests that assess memory capacity. However, on even more poorly *g* loaded and thus easier measures than the K-ABC memory subtests, white performance may actually be crippled because the low anxiety white profile may not produce the arousal, attentional, and motivational levels (i.e., an “increase in drive level”) needed for successful completion of these relatively cognitively superficial tasks. Indeed, recall that Rocklin and Thompson’s low anxiety subjects performed better on a hard test version than on an easy version. In summary, highly *g* loaded tests may bias against blacks, but poorly *g* loaded tests may bias against whites.

Concluding Remarks: Toward a More Perfect Cognitive Ability Assessment

So where should the future of intelligence testing lie? Certainly our examination of race-dependent differential socialization experiences emphasizes the need to consider the complex underlying bases of the black-white performance difference. We have seen that classical intelligence batteries and the K-ABC are both imperfect measures. However, they are far from being useless psychometric tools. They are just instruments that must be used with extreme caution. A simple IQ score proffered by these tests only reflects the cognitive capacity of a black or white examinee to a first approximation. The examinee’s racially determined socialization experiences must be factored into the assessment of mental ability. Although such a qualitative individualized appraisal is laborious, it may be the only equitable assessment procedure.

To summarize my socialization-mediated self-defeating model, I have argued that black children's lack of a maternally created zone of proximal development and consequent cognitive rigidity may preclude success on highly *g* loaded tasks. Data on black children's spontaneous elaboration behavior, the rather authoritarian matriarchal organization of the black household, black maternal use of aversive discipline, negative tension release, and negative evaluations of the child's mental task efforts, and black stereotype anxiety were then elicited to demonstrate black children's high test anxiety. This high anxiety may especially interfere with black children's test performance subsequent to probable failure on highly *g* loaded items due to such factors as the stable internal causal attributions that high anxiety blacks adopt for their failure. In turn, subsequent persistence is compromised and exit behaviors are espoused by blacks to disengage from test-taking (recall the data on exit not-work responses). This synopsis of a highly complex chain of events highlights the central mediating role that socialization experiences undoubtedly play in governing intellectual test performance.

The goal of achieving more racially impartial cognitive assessment methods can only be achieved by viewing the poorer performance of blacks as the product of differences in such domains as socialization experiences rather than due to deficits possessed by black children. Many psychologists have already begun to adopt such an evaluative framework.⁴³⁻⁴⁶ Only when this more qualitative socialization analysis is combined with quantitative performance on the WISC-R, Stanford-Binet, or K-ABC will a more "enlightened" assessment of a child's cognitive ability have been achieved. It is imperative that psychologists adopt such global evaluative methods, for otherwise the mental capacities of certain children may be severely underestimated. Because the future of so many children is at stake, there can be little room for error.

Acknowledgments

I would like to thank Ronald Butzlaff for all the guidance he has provided me in the construction of this paper and for his recognizing that a social conscience must pervade the products of academia. I would also like to thank Stephen M. Kosslyn, Ph.D. and Bruno Laeng, Ph.D. for their inspiration to discover the joys of studying theoretical and empirical psychology.

References

- (1) Gupta, R. M., and P. Coxhead (eds.). 1988. *Cultural diversity and learning efficiency: recent developments in assessment* (New York: St. Martin's Press).
- (2) Hickman, J. A., and C. R. Reynolds. 1986-87. "Are race differences in mental test scores an artifact of psychometric methods? A test of Harrington's experimental model." *The Journal of Special Education*. 20(4): 409-30.
- (3) Jensen, A. R. 1984. "The black-white difference on the K-ABC: implications for future tests." *The Journal of Special Education*. 18(3): 377-408.
- (4) Gaudry, E., and D. Fitzgerald. 1971. "Test anxiety, intelligence and academic achievement." In *Anxiety and educational achievement* (eds. Gaudry, E., and C.D. Spielberger) (New York: John Wiley & Sons Australasia Pty).
- (5) Van der Ploeg, H. M., R. E. Hulshof. 1984. "Cross-cultural study of the relationship among academic performance, test anxiety, intelligence, and sex." *Psychological Reports*. 55: 343-6.
- (6) Zigler, E., and E. C. Butterfield. 1968. "Motivational aspects of changes in IQ test performance of culturally deprived nursery school children." *Child Development*. 39: 1-14.
- (7) Terman, L. M., M. Merrill. 1960. *Stanford-Binet intelligence scale: manual for the third revision. Form L-M* (Boston: Houghton-Mifflin).
- (8) Wechsler, D. 1974. *Manual for the Wechsler Intelligence Scale for Children-Revised* (New York: The Psychological Corporation).
- (9) Peoples, C. E., J. F. Fagan, III, and D. Drotar. 1995. "The influence of race on 3-year-old children's performance on the Stanford-Binet: Fourth Edition." *Intelligence*. 21: 69-82.
- (10) Montie, J. E., and J. F. Fagan, III. 1988. "Racial differences in IQ: item analysis of the Stanford-Binet at 3 years." *Intelligence*. 12: 315-32.
- (11) Naglieri, J. A., and A. R. Jensen. 1987. "Comparison of black-white differences on the WISC-R and the K-ABC: Spearman's hypothesis." *Intelligence*. 11: 21-43.
- (12) Kaufman, A. S., and N. L. Kaufman. 1983. *K-ABC: Interpretive manual* (Circle Pines, MN: American Guidance Service).
- (13) Anastasi, A. 1988. *Psychological testing*, 6th ed. (New York: McMillan).
- (14) Jensen, A. R. 1980. *Bias in mental testing* (New York: The Free Press).
- (15) Jensen, A. R. 1983. "Do IQ tests really measure intelligence?" *The Behavioral and Brain Sciences*. 6: 311-5.
- (16) Herrnstein, R. J., and C. Murray. 1994. *The bell curve: intelligence and class structure in American life* (New York: The Free Press).
- (17) Detterman, D. K., and M. H. Daniel. 1989. "Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups." *Intelligence*. 13: 349-59.
- (18) Spearman, C. 1927. *The abilities of man: their nature and measurement* (New York: Macmillan).
- (19) Moore, E. G. J. 1986. "Family socialization and the IQ test performance of traditionally and transracially adopted black children." *Developmental Psychology*. 22(3): 317-26.
- (20) Scarr, S., and R. A. Weinberg. 1976. "IQ test performance of black children adopted by white families." *American Psychologist*. 31: 726-39.
- (21) Weinberg, R. A., S. Scarr, and I. D. Waldman. 1992. "The Minnesota transracial adoption study: a follow-up of IQ test performance at adolescence." *Intelligence*. 16: 117-35.
- (22) Turkheimer, E. 1991. "Individual and group differences in adoption studies of IQ." *Psychological Bulletin*. 110(3): 392-405.
- (23) Cole, M., and S. R. Cole. 1993. *The Development of Children*, 2nd ed. (New York: W.H. Freeman and Company).
- (24) Hess, R. D., and V. C. Shipman. 1967. "Cognitive elements in maternal behavior." In *Minnesota symposia on child psychology*, vol. 1 (ed. Hill, J. P.) (Minneapolis: The University of Minnesota Press).
- (25) Curtis, R. C. (ed.). 1989. *Self-defeating behaviors: experimental research, clinical impressions, and practical implications* (New York: Plenum Press).
- (26) Miele, F. 1979. "Cultural bias in the WISC." *Intelligence*. 3: 149-64.
- (27) Payne, B. D., J. E. Smith, and D. A. Payne. 1983. "Grade, sex, and race differences in test anxiety." *Psychological Reports*. 53: 291-4.

- (28) Jerusalem, M., D. Liepmann, and C. Herrmann. 1987. "Test anxiety and achievement motivation: an analysis of causal relationships." In *Advances in Test Anxiety Research* (eds. Schwarzer, R., H. M. Van der Ploeg, and C. D. Spielberger) (Berwyn, PA: Swets North America).
- (29) Zigler, E., W. D. Abelson, and V. Seitz. 1973. "Motivational factors in the performance of economically disadvantaged children on the Peabody Picture Vocabulary Test." *Child Development*. 44: 294-303.
- (30) Seitz, V., W. D. Abelson, E. Levine, et al. 1975. "Effects of place of testing on the Peabody Picture Vocabulary Test scores of disadvantaged Head Start and non-Head Start children." *Child Development*. 46: 481-6.
- (31) Dusek, J.B. 1980. "The development of test anxiety in children." In *Test Anxiety: Theory, Research, and Applications* (ed. Sarason, I. G.) (Hillsdale, NJ: Lawrence Erlbaum Associates).
- (32) Crockenberg, S. 1983. "Early mother and infant antecedents of Bayley Scale performance at 21 months." *Developmental Psychology*. 19(5): 727-30.
- (33) Blau, Z. S. 1981. *Black children/white children: competence, socialization, and social structure* (New York: The Free Press).
- (34) Gay, G., and R. D. Abrahams. 1973. "Does the pot melt, boil, or brew? Black children and white assessment procedures." *Journal of School Psychology*. 11(4): 330-40.
- (35) Steele, C. M. 1995. "Stereotype threat and the intellectual test performance of African Americans." *Journal of Personality and Social Psychology*. 69(5): 797-811.
- (36) Leppin, A., R. Schwarzer, D. Belz, et al. 1987. "Causal attribution patterns of high and low test-anxious students." In *Advances in Test Anxiety Research* (eds. Schwarzer, R., H. M. Van der Ploeg, and C. D. Spielberger) (Berwyn, PA: Swets North America).
- (37) Hedl, J. J., Jr. 1987. "Explorations in test anxiety and attribution theory." In *Advances in Test Anxiety Research* (eds. Schwarzer, R., H. M. Van der Ploeg, and C. D. Spielberger) (Berwyn, PA: Swets North America).
- (38) Covington, M. V., and C. L. Omelich. 1987. "Item difficulty and test performance among high-anxious and low-anxious students." In *Advances in Test Anxiety Research* (eds. Schwarzer, R., H. M. Van der Ploeg, and C. D. Spielberger) (Berwyn, PA: Swets North America).
- (39) Rocklin, T., and J. M. Thompson. 1985. "Interactive effects of test anxiety, test difficulty, and feedback." *Journal of Educational Psychology*. 77(3): 368-72.
- (40) Arkin, R. M., and Haugtvedt. 1987. "Test-anxiety, task difficulty and diagnosticity: the roles of cognitive interference and fear of failure as determinants of choice of task." In *Advances in Test Anxiety Research* (eds. Schwarzer, R., H. M. Van der Ploeg, and C. D. Spielberger) (Berwyn, PA: Swets North America).
- (41) Covington, M. V. 1987. "Anxiety, task difficulty and childhood problem-solving: a self-worth interpretation." In *Advances in Test Anxiety Research* (eds. Schwarzer, R., H. M. Van der Ploeg, and C. D. Spielberger) (Berwyn, PA: Swets North America).
- (42) Schwarzer, R., M. Jerusalem, and A. Stiksrud. 1987. "The developmental relationship between test anxiety and helplessness." In *Advances in Test Anxiety Research* (eds. Schwarzer, R., H. M. Van der Ploeg, and C. D.) (Berwyn, PA: Swets North America).
- (43) Cole, M., and J. S. Bruner. 1971. "Cultural differences and inferences about psychological processes." *American Psychologist*. 26: 867-76.
- (44) Yeates, K. O., D. MacPhee, F. A. Campbell, et al. 1983. "Maternal IQ and home environment as determinants of early childhood intellectual competence: a developmental analysis." *Developmental Psychology*. 19(5): 731-9.
- (45) Clark, L. A., and G. S. Halford. 1983. "Does cognitive style account for cultural differences in scholastic achievement?" *Journal of Cross-Cultural Psychology*, 14(3): 279-96.
- (46) Borkowski, J. G., and A. Krause. 1983. "Racial differences in intelligence: the importance of the executive system." *Intelligence*. 7: 379-95.

