The nature of nurture: the role of gene-environment interplay in the development of intelligence

Kan, K.-J.

Publication date
2012

CHAPTER 4

A GENETIC ORIGIN OF BLACK-WHITE MEAN IQ DIFFERENCES? WEAK INFERENCES BASED ON AMBIGUOUS RESULTS

Abstract

Recently, Intelligence published an editorial with the aim to clarify the relation between subtests' g *loadings, heritability coefficients, and Black-White mean IQ differences. The following conclusions were drawn: (1) There is no strictly nongenetic (e.g., environmental or cultural) explanation for positive correlations among* g *loadings, heritability coefficients, inbred-outbred differences, and mean Black-White differences; and (2) These positive correlations support the 'biological* g *theory'. We demonstrate that the reasoning underpinning the first conclusion is invalid. The correlations have no bearing on the relative contributions of genetic and environmental effects on the mean Black-White differences. In addition, we scrutinize Black-White IQ data in the light of biological* g *theory. This theory predicted that the mean differences are larger for culture reduced tests than culture loaded tests. The data show the opposite*

4.1 A Genetic Origin of Black-White Mean IQ Differences?

Group differences in intelligence, racial differences in particular, are the topic of extensive and often heated debate. In this debate, political opinions and emotions often overrule scientific arguments. However, scientific research into these differences is legitimate, and empirical results can lead to the conclusion that racial differences in mean IQ, are, like individual differences, to some extent due to genetic differences. One complication within this research is that mean group differences can be entirely of environmental origin, while individual differences can be highly heritable (Lewontin, 1970; Rushton & Jensen, 2005). To date, the relative contributions of genetic and environmental influences to phenotypic IQ group differences remain generally unknown. In a recent editorial in *Intelligence*, however, Rushton and Jensen (2010a) maintained that certain empirical results help to reveal the origin of group differences. Such results are thus potentially important in the scientific debate on the origins of racial differences.

With their editorial Rushton and Jensen (2010a, henceforth R&J) set out to "clarify the relation between *g* loadings, heritabilities, Black-White differences, and the secular rise in IQ." (p. 214). In effect, they reviewed intercorrelations among these and a number of other variables, including the correlation between Black-White mean IQ differences and inbreeding effects (see R&J, p. 216, Table 1). From these correlations R&J drew various conclusions concerning the origin (environmental or genetic) of IQ group differences in general, and of Black-White IQ mean differences and secular gains in particular. In the light of these conclusions, R&J discussed implications for competing theories and explanations of Black-White differences. Although we are willing to entertain the hypothesis that mean Black-White differences (or indeed any phenotypic mean differences in valid measures) are partly genetic in origin, we demonstrate that the reasoning underpinning some of R&J's conclusions is invalid. Additionally, we demonstrate, by means of a re-analyses of R&J's data, that Black-White mean differences are inconsistent with the theory of intelligence that R&J favor.

The conclusions in the R&J's editorial can be summarized as follows: (1) the general factor of intelligence represents a real - that is, biological - largely genetically influenced variable (*g*), (2) the omnipresent secular gains in IQ are of environmental origin, and, (3) do not reflect differences (gains) on this biological variable, while (4) mean Black-White differences in IQ do reflect differences on this biological variable and (5) are, in part, of genetic origin. Conclusion 1 was based on the finding that IQ subtests' *g* loadings correlate positively with their heritability coefficients (h^2).

Conclusion 2 was based on the finding that the subtests secular gains are zero or negatively correlated with their heritability coefficients. Conclusion 3 was based on the finding that these gains are zero or negatively correlated with *g* loadings. Conclusions 4 and 5 were based on the finding that Blacks score on average lower than Whites, and the mean differences are largest on the most *g* loaded subtests (see, e.g. Jensen 1885; 1987), which are also the most heritable subtests, and on the finding that the (standardized) mean Black-White differences correlate positively with the heritability coefficients themselves.

R&J considered results pertaining to samples that are characterized by a degree of inbreeding depression to gather support for their conclusion that the Black-White mean differences are partly of genetic origin. They offered the following rationale (see R&J, p. 214). First, inbreeding depression is 'a purely genetic effect'. Second, because inbreeding depressed samples score systematically lower on IQ tests than outbred samples, and the subtests' *g* loadings and heritabilities correlate positively with the mean differences, inbreeding depression influences IQ via *g*. Third, the mean Black-White differences correlate positively with inbred-outbred differences. The upshot was that 'there is no non-genetic explanation' (R&J, p. 214) for the finding that *g* loadings, heritabilities, inbreeding effects, and mean Black-White differences intercorrelate positively.

One of the purposes of R&J was to provide empirical evidence that, in contrast to the subtests' secular gains, which are generally considered to be of environmental origin (but see Mingroni, 2007), Black-White mean differences show a pattern similar to the inbred-outbred mean differences (considered to be a purely genetic effect), and are, at least in part, of genetic origin. That is, R&J reviewed evidence that mean Black-White differences are the largest on the most heritable, most *g* loaded subtests. This purpose appears to make sense in light of their hypothesis that "if population group differences are greater on the more *g* loaded and more heritable subtests, it implies they have a genetic origin." (R&J, p. 214). From this hypothesis, which we will denote the genetic origin hypothesis, R&J inferred the following (R&J, p. 214):

> "(1) Genetic theory predicts a positive association between heritability and group differences; (2) culture theory predicts a positive association between environmentality and group differences; (3) nature+nurture models predict both genetic and environmental contributions to group differences; while (4) culture-only theories predict a zero relationship between heritability and group differences." (Rushton & Jensen, 2010a, p. 214)

Our present aims are threefold. First, we show that one cannot infer that these group differences are genetic because these differences are greater on the more *g* loaded and more heritable subtests. Secondly, we demonstrate that the predictions of the competing theories, as conveyed by R&J, are open to debate. Specifically, contrary to R&J's inference concerning culture theory and culture only theory, these theories can readily accommodate a positive correlation between heritability and *g* loading. Thirdly, in the light of Rushton and Jensen's biological *g* theory (Rushton & Jensen, 2005; 2010a, 2010b), we reexamine the original Black-White data in Jensen (1985, 1987), to which R&J refer. We show that one important association remains to be clarified, namely the relation between the subtests' Black-White mean differences and their cultural loading. We conclude that 1) the relative contributions of genes and environment to Black-White differences remain unknown and 2) biological *g* theory does not give a sufficient explanation of the relation between cultural loading and Black-White differences.

4.2 The Reasoning Underlying the Genetic Origin Hypothesis is Weak

We first demonstrate that the reasoning underlying the genetic origin hypothesis is open to criticism because it involves an *affirmation of the consequence*. Consider the (standardized) regression model in Figure 4.1, in which two groups are compared with respect to IQ. Within each group, the IQ subtest scores are regressed on the (latent, unmeasured) variable *g* (with regression weights λ_i); *g*, in turn, is regressed on latent (unmeasured) genetic and environmental influences (with regression weights *h* and *e*, respectively). We introduce a mean group difference on the genetic component of *g*. The effect is purely genetic, as we assume that the means of the environmental components over the groups are equal. The genetic mean difference results in a group mean difference in *g*. Because there is a mean difference in *g*, the IQ mean differences are the most pronounced on the subtests with the largest the *g* loadings (i.e., the subtests' with the largest regression weights on *g*). The *g* loadings (λ_i) correlate perfectly with the subtests' regression weights on *g*'s genetic component, which take the value of $h\lambda_i$. Because in this model, *g* is the only source of the subtests genetic variance, it follows that the subtests' heritability coefficients take the value of the squared regression weights on *g*'s genetic component: $(h\lambda_i)^2$. The relation between the subtests' *g* loadings (λ_i) and their heritability coefficients ($h^2\lambda_i^2$) is thus also perfect (although quadratic). Because both the IQ mean differences and the regression weights on *g*'s genetic component ($h\lambda_i$) are collinear with the *g* loadings (λ_i), the (rank) correlation between the subtests' IQ mean differences and their heritability coefficients ($h^2\lambda_i^2$) is +1.

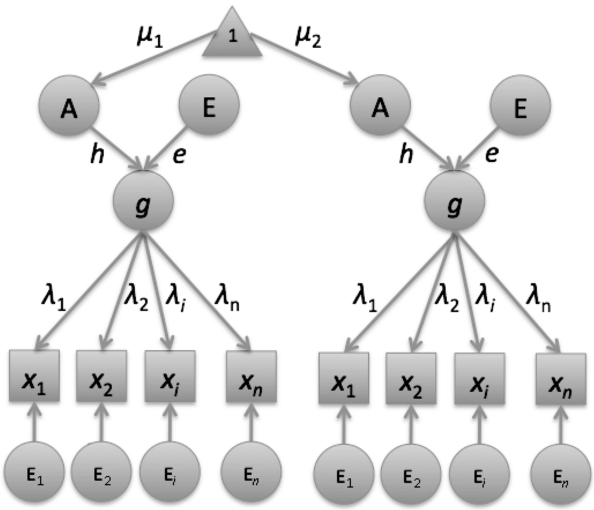

Figure 4.1 A mean difference on the genetic component of *g* will result in a mean difference on *g*. A mean difference in *g*, in turn, will result in mean differences in IQ. The ranks of the subtests' mean differences will correlate perfectly with the ranks of the subtests' heritability coefficients (which take the value of $h^2\lambda_i^2$), because both correlate perfectly with the ranks of the subtests' *g* loadings.

Thus, due to a purely genetic origin, the population group differences are greater on the more *g* loaded and more heritable subtests. We let the two groups represent an inbreeding depressed sample and an outbred white sample (we assume that the inbreeding depressed sample has a lower average level of *g*: $\eta_1 < \eta_2$).

Next, we compare the normal white sample with a normal black sample (Figure 4.2). We introduce a mean difference on *g* (but not on the more specific factors of intelligence), such that blacks have a lower average than whites ($\eta_1 < \eta_2$). Like the inbred-outbred IQ mean differences, the Black-White IQ mean differences will be the most pronounced on the most *g* loaded tests (those with the highest λ_i). So, these differences will correlate perfectly with the inbred-outbred mean differences. Because *g* loadings (λ_i) are collinear with the regression weights on *g*'s genetic component ($h\lambda_i$), the Black-White gaps are also collinear with them. As a corollary, the rank correlation between Black-White gaps and heritability coefficients ($h^2\lambda_i^2$) is +1. In short, the intercorrelations among inbred-outbred mean differences, Black-White mean differences, heritability coefficients, and *g* loadings are positive and perfect.

In view of the genetic origin hypothesis, the crucial question is: Does this finding *imply* that the Black-White differences has a genetic origin? Is there 'no non-genetic explanation' for the finding that *g* loadings, heritabilities, inbreeding effects and mean Black-White differences intercorrelate positively?

We agree that it is *possible* that the mean Black-White difference in *g* is due in part to a mean difference on *g*'s genetic component (as in Figure 4.1). If so, in studying Black-White differences, we would expect find the pattern of correlations among *g* loadings, heritabilities, and mean group differences similar to those found in inbred-outbred studies. However, inferring an underlying cause from this pattern of correlations is problematic in view of the following.

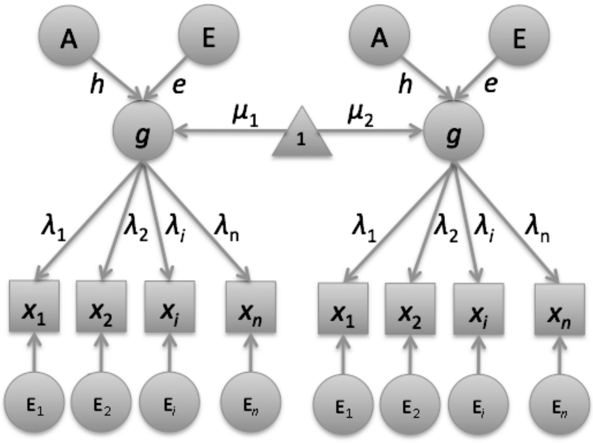

Figure 4.2 A mean difference in *g* will result in mean differences in IQ. The ranks of the subtests' mean differences will correlate perfectly with the ranks of the subtests' heritability coefficients (which take the value of $h^2\lambda_i^2$). The source (genetic or environmental) of this mean difference is unknown.

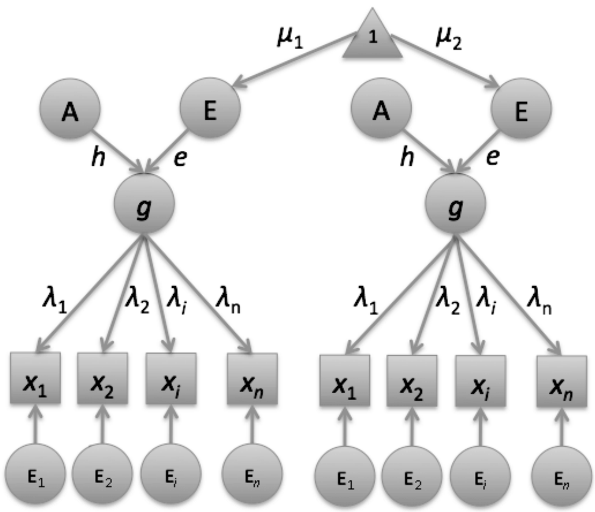

Figure 4.3 A mean difference on the environmental component of *g* will result in a mean difference on *g*. A mean difference in *g*, in turn, will result in mean differences in IQ. The ranks of the subtests' mean differences will correlate perfectly with the ranks of the subtests' heritability coefficients (which take the value of $h^2\lambda_i^2$), because both correlate perfectly with the ranks of the subtests' *g* loadings (λ_i).

Consider the situation depicted in Figure 4.3. That is, suppose that the true situation was that Blacks and Whites differed on *g*'s *environmental* component. As a result, there would be a mean difference on *g*. Again, the mean IQ differences would be greatest on the most *g* loaded tests (i.e., those with the highest λ_i), which happen to be most heritable tests (as shown above, the subtests' heritability coefficients will equal $h^2\lambda_i^2$). The Black-White mean IQ differences would correlate perfectly with the inbred-outbred mean IQ differences, since they are both collinear with *g* loading (λ_i). Thus, in this situation, in which the Black-White differences were purely of environmental origin, we would have found the same pattern of correlations among Black-White IQ differences, inbred-outbred IQ differences, *g* loadings, and heritability coefficients as in the previous situation, in which the origin of the group differences was genetic. So, in this *g* model, any relative contribution of genes and environment to group differences could lead to the finding that Black-White differences, inbred-outbred differences, *g* loadings, and heritabilities intercorrelate positively. Correlations among these variables have no necessary bearing on the origin of group differences, and do not furnish a basis for inferences, strong or otherwise.

4.3 Inferences are Invalid

We now examine R&J's (p. 214) inferences concerning 'genetic theory' and 'culture (only) theory'. As explained in Rushton & Jensen (2005), genetic theory "contends that a substantial part (say 50%) of both individual and group differences in human behavioral traits is genetic" (Rushton & Jensen, 2005, p. 238), while culture only theory "finds no need to posit any genetic causation" (Rushton & Jensen, 2005, p. 238) of group differences:

> "*The defining difference* [between genetic theory and culture-only theory] *is whether any significant part of the mean Black-White IQ difference is genetic rather than purely cultural or environmental in origin*." (Rushton & Jensen, 2005, p. 238, italics original)

In the light of this distinction between 'genetic theory' and 'culture-only theory, consider the first inference: "Genetic theory predicts a positive association between heritability and group differences." (R&J, p. 214). On the one hand, it is true that genetic theory can predict this association, but under certain circumstances. It is the case, for example, if *g* exists and if *g* is the sole (or main) source of subtests genetic variance, and if there is a group difference on *g*'s genetic component. On the other hand, as demonstrated above, a (possibly perfect) positive correlation between heritability and group differences is quite compatible with purely environmental origins of the phenotypic mean Black-White differences. That 'culture-only theories predict a zero relationship between heritability and group differences' is thus not necessarily true. In addition, in more complex models than the simple 1-factor model with the common factor as the only source of subtests' genetic variance, mean differences can be entirely genetic without a positive association between subtests' heritabilities and mean differences. One can only make unambiguous predictions about the correlation between heritabilities and mean group differences, if one explicitly models the causal genetic and environmental pathways, for example in a Structural Equation Model.

The third inference that "culture theory predicts a positive association between environmentality and group differences" (R&J, p.214) is not also necessarily true. In Figure 4.3, for example, the origin of mean group differences is purely environmental, but the correlation between group mean differences and environmentalities (1- $h^2\lambda_i^2$) is -1. Analogous to the argument above, one cannot make predictions about the correlation between environmentalities and mean group differences without explicitly modeling the causal pathways.

We conclude that nothing about genetic and culture-only theories can be *inferred* from the vector correlations. Thus, in order to establish with any certainty the pathways through which genetic and environmental latent variables affect the factors of intelligence, these variables and pathways must be modeled explicitly. This is not sufficient to establish the contributions of environmental and genetic sources of the mean group differences, however, because in the latent *g* model these contributions are unidentified. To establish these, environmental and genetic variables should have been measured and incorporated in the model, replacing in part or entirely the latent variables. With respect to genetic influences, this requires the identification of genes that explain (preferably in a causal sense) individual differences in intelligence. Once identified, their actual contribution to between group, i.e., Black-White, differences can be estimated. The same applies to environmental (causal) influences.

Besides the fact that the inferences are not valid, R&J's explanation of Black and White IQ mean differences is problematic. Specifically, as we discuss in the next section, the data are not consistent with respect to R&J's position on the role of cultural loading.

4.4 Problems for Biological *g* Theory

When comparing competing theories of intelligence or explanations of Black-White differences, one needs falsifiable predictions. Here, we argue that in addition to the problem with Rushton and Jensen's inferences (as identified above), predictions stemming from their biological *g* theory are not consistent with certain empirical results concerning Black-White mean differences. That is, if we accept - for the sake of argument - Rushton and Jensen's method (essentially Jensen's method of correlated vectors; Jensen, 1998, pp. 372-374), and assuming that Black-White mean differences are in indeed due to the biological variable *g* and partly of genetic origin, we are confronted by puzzling empirical results concerning the role of cultural loading in the Black-White differences.

In biological *g* theory, the variable *g* is viewed as highly heritable and largely fixed, i.e., not readily malleable by cultural influences. Here cultural influences are contrasted with genetic influences, and correspondingly cultural loadings are contrasted with *g* loadings (see Rushton & Jensen, 2005, 2010b; Rushton, 1998). Culture reduced test are hypothesized to have higher *g* loadings and higher heritability coefficients than culture loaded tests. Hence, the hypothesis that Black and White group differences are more pronounced on the more *g* loaded tests, implies that in biological *g* theory the Black-White mean differences are relatively small on strongly cultural loaded tests ('culture loaded' tests) and relatively large on weakly cultural loaded tests ('culture reduced' tests). Rushton & Jensen (2005, p.272) expressed this as follows: "Mean Black-White differences [are] greater on *g* loaded cognitive tests than on culturally loaded cognitive tests".

Culture theory, as conveyed by R&J, on the other hand, "predicts that differences between races will be greater on those culturally malleable items on which races can grow apart as a result of dissimilar experiences." (Rushton, 1998, p. 222). We note that 'culture theory' differs from 'culture only theory' (Rushton & Jensen, 2005) in the sense that culture theory does not necessarily exclude the possibility that genetic differences play a role in the development of individual and group differences. Consider the theory of Dickens & Flynn, 2001), for example. In this theory intelligence is (highly) heritable, but also malleable by environmental (e.g. cultural) influences. These influences can increase initial (individual and mean group) differences, whether these are genetic or environmental. A culture theory can thus encompass a genetic theory (a theory that describes that a substantial part of both individual and group differences in intelligence is genetic), or a 'nature-nurture model' (R&J, p. 214).

Both biological *g* theory and Dickens' and Flynn's theory are nature-nurture models, but the former predicts that Black-White differences are greater on culture reduced tests than culture loaded test, while the latter predicts that these differences are greater on cultural reduced tests than culture loaded tests. In the following section, we first explain the concept of cultural loading in some more detail. Next, using data upon which R&J based their conclusions, we evaluate the relation between Black-White differences, *g* loading and cultural loading in the light of biological *g* theory.

4.4.1 Cultural loading

Jensen provided the following conceptualization of a test's cultural loading:

> "Tests and test items can be ordered along a continuum of culture loading, which is the specificity or generality of the informational content of the test items. The narrower or less general the culture in which the test's information could be acquired, the more cultural loaded it is. A test may contain information that could only be acquired within a particular culture. This can actually be determined simply by examination of the test items. The specificity or generality of the content corresponds to its cultural loading." (Jensen, 1976, p.340)

Culture loaded tests are thus the tests that occupy relatively high positions on the cultural loading continuum. At the other end of the continuum are located 'culture reduced' (or 'culture fair') items or tests. The theoretical distinction between culture loaded and culture reduced tests features quite often in the literature (e.g. Hunt & Sternberg, 2005; te Nijenhuis, Tolboom, Resing, & Bleichrodt, 2004; Helms-Lorenz & van de Vijver, 2003), and different operationalizations of cultural loading have been used (e.g. te Nijenhuis et al., 2004; Helms-Lorenz & van de Vijver, 2003). One operationalization involves establishing loading on the basis of ratings by experts or ratings by students, another involves the categorization of tests into 'aptitude tests' and 'achievement tests'. These operationalizations are consistent with each other: Knowledge, scholastic, and achievement test are viewed as most cultural loaded (see also Chapter 3).

In this connection, the distinction between crystallized and fluid abilities is also relevant. Specifically, we note that some researchers equate culture loaded tests with 'crystallized abilities' tests (e.g. Hunt & Sternberg, 2005). Indeed, crystallized abilities refer to knowledge and well-practiced skills, and hence are highly culturally influenced (Cattell, 1987). They are differentiated from 'fluid abilities', which concern reasoning abilities. Although the process of solving the items on fluid abilities tests always requires knowledge to some degree, hence is culturally influenced to some degree, this knowledge is considered to be equally known or equally new for the people for whom the test is valid. This implies that, in contrast to crystallized ability tests, prior knowledge is assumed not to *differentiate* between the members of the population of interest, and is therefore no source of individual differences.

The distinction between fluid and crystallized abilities tests shows a large overlap with the distinction between culture loaded versus culture reduced tests (see Chapter 3), but they do differ (Cattell, 1987; Jensen, 1998). Verbal *knowledge* tasks, for example, are crystallized ability tests, while tests of verbal *reasoning* are fluid ability tests. However, the assumption that the required knowledge in verbal reasoning is equally known is reasonably tenable only if the testees share the same linguistic and cultural background. The items of verbal reasoning tests have to be translated or adjusted before they can be administered to testees from other cultural backgrounds; otherwise the tests would be invalid. Both verbal (crystallized) knowledge and verbal (fluid) reasoning tests are culture loaded, but the latter less than the former.

Here, we prefer to retain the terms culture reduced and culture loaded, because the terms fluid and crystallized abilities are linked to a particular theory of intelligence (Cattell's investment theory of fluid and crystallized intelligence; e.g. Cattell, 1987), whereas the concept of cultural loading is more independent of theories of intelligence.

Henceforth, in our statistical analysis, we abide by Jensen's definition of culture reduced tests, and dichotomize cultural loading in the light of this definition: Culture reduced tests are "those that are nonlanguage and nonscholastic and do not call for any specific prior information for a plus-scored response" (Jensen, 1980, p. 374). Therefore, we consider tests that depend on language, that are scholastic or that require specific substantive prior information as culture loaded (Jensen, 1980, p.). Examples of culture loaded tests are the (WISC and WAIS) subtests Vocabulary, Information, and Similarities (verbal knowledge tests, calling for specific knowledge), Verbal Comprehension (a verbal reasoning test, depending on language) and Arithmetic (a scholastic test) (see also Chapter 3). Examples of tests that are considered to be culture reduced are Raven's Progressive Matrices (nonverbal, abstract reasoning), Block Design (mental manipulation in space), and elementary cognitive tasks (e.g. reaction time tasks) (see also Chapter 3).

4.4.2 Reanalyses of the Data: Black-White Differences are Largest on the Most Culturally Loaded Tests

Predictions Rushton & Jensen (2005; 2010b) implied that *g* loading and cultural loading are intrinsically inversely, i.e., negatively, related. As quoted above, on the basis of biological *g* theory, Rushton and Jensen hypothesized (and also reported, e.g. Jensen, 1985; Rushton, 1995) that mean Black-White differences in IQ are larger on *g* loaded than on culture loaded tests. One would expect the data to be in line with Figure 4.4.

We note that contrary to the biological *g* prediction, but compatible with Dickens and Flynn's theory (2001; Dickens, 2008), cultural loading and *g* loading may well be intrinsically *positively* related (see Chapter 3), for example when *g* loadings are a function of societal demands (Dickens, 2008). The finding that Black-White differences are largest on the most *g* loaded, most heritable tests, poses no problem for this theory.

We argue that empirical data are not in line with the predictions from biological *g* theory. To this end, we present the results of a re-analyses of data, to which R&J (Jensen, 1998, p. 378, 386) refer.

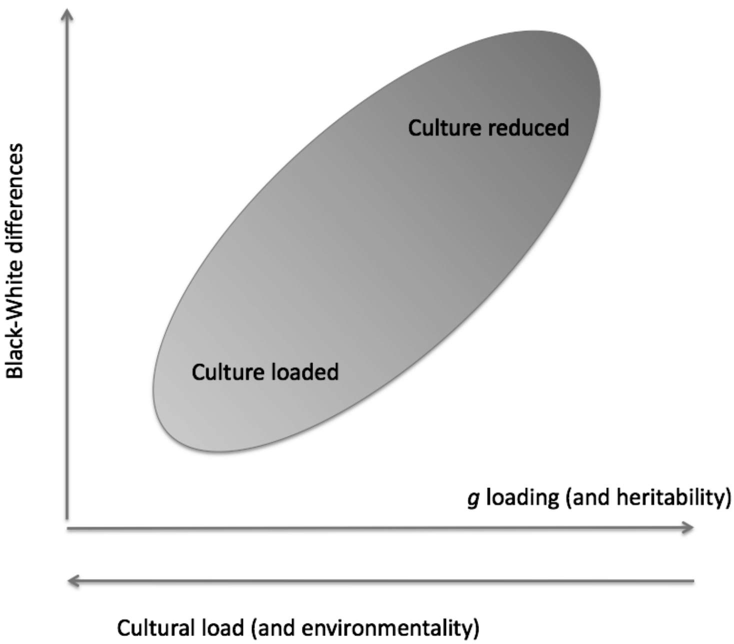

Figure 4.4 In biological *g* theory, *g* is viewed as highly heritable and not readily malleable by cultural influences. Cultural influences are contrasted with genetic influences; cultural loading with *g* loading. Culture reduced tests have higher *g* loadings and higher heritability coefficients than culture loaded tests. Black-White mean differences are relatively small on culture loaded tests and relatively large on culture reduced tests.

Data The data are summarized in a review of Black-White differences by Jensen (1985, 1987), who intended to show that the standardized mean differences are a function of *g* loading. They come from 11 data sets in Jensen's (1985) target article in Behavioral and Brain Sciences (1985), together with the additional single dataset from a later commentary on this article (Jensen, 1987). In these studies, the loadings on the first principal factor featured as the *g* loadings.

Participants The total set contained 145 Black and White standardized mean differences and *g* loadings. Study characteristics and data are provided in these papers. The subject samples were representative for the US Black and White populations (e.g. subjects participated in national studies or in standardization procedures). The total number of participants was 26464 (1560 whites and 10824 blacks).

Method The total number of unique subtests was 73 (see Table 4.1 for details). Within each study, we averaged subtests' *g* loadings in the black and the white samples (we weighted by sample size of blacks and whites). In most cases these were given separately. If not, we assumed *g* loadings were equal in both groups. We considered this acceptable because the comparison of factor loadings revealed that they were similar (Jensen, 1985). For each unique subtest, we calculated the weighted means of *g* loading, reliability, and Black-White differences across studies (we weighted by the studies' total sample size).

Subtests' cultural load was determined by examining the subtests descriptions or the description of the factors on which they loaded. If the subtests measured scholastic abilities, achievement, crystallized intelligence, verbal comprehension, arithmetic, knowledge (or synonyms of these terms), we coded cultural load as 1, indicating that these tests are 'culture loaded' tests. The cultural load of the other tests was coded 0 to absence of appreciable cultural load. The proportion culture loaded test so defined was approximately 53%.

Results Visually inspecting the original data we noticed that within each battery the largest Black-White standardized mean differences were on culture loaded tests (most often involving vocabulary; see Jensen 1985, p.203). Also in the complete dataset (see the scatterplot in Figure 4.5), the Black-White mean differences appeared the largest on culture loaded tests.

The boxplot in Figure 4.6 (left pane) shows that the subtests' Black-White differences were indeed larger on culture loaded tests than on culture reduced tests. This is inconsistent with biological *g* theory. We also note that Rushton and Jensen's culture loaded *versus* *g* loaded distinction is problematic, because *g* loadings were the highest for culture loaded tests (see the boxplot in Figure 4.6, middle pane, see also Chapter 3). The intercorrelations between the subtests' cultural load, averaged *g* loadings, and averaged Black-White differences were all positive and significant at $\alpha = .01$.

However, as Jensen (1987) mentioned, reliabilities of culture loaded tests were higher than reliabilities of culture reduced tests (see the boxplot in Figure 4.6, right pane), which have resulted in underestimation of factor loadings of culture reduced test. Below, we show that even if we accept, for the sake argument, R&J Jensen's method of correlated vectors (p. 372-374), we would arrive at the same conclusion: Black-White differences are largest on culture loaded tests.

The method prescribes that reliability should be taken into account out by adjusting *g* loadings and Black-White standardized mean differences using correction for attenuation. We considered biological *g* theory not to be in conflict with the data when if this is done, *g* loading correlates with mean Black-White differences, while *g* loading and cultural load do not correlate positively. We corrected *g* loadings and Black-White differences for attenuation. That is, we divided them by the square root of their reliabilities. Next, we calculated the intercorrelations among cultural load, the corrected *g* loadings and the corrected Black-White differences. These were all positive (see Table 4.2) and significant at a significant level of $\alpha = .01$. Culture loaded tests still showed larger Black-White differences (mean = 0.78) than culture reduced tests (mean = 0.63) ($t = -5.84$, df = 57, $p < 0.001$).

Conclusion Black-White mean differences largest on the most culture loaded subtests whereas cultural load and *g* loading correlate positively. These findings are problematic for biological *g* theory, because in this theory cultural load and *g* loading are juxtaposed; it was predicted that Black-White mean differences are largest on cultural reduced subtests. We maintain that we need more adequate theory to explain these data. As the contributions of genes and environment to the Black-White mean IQ differences cannot be determined or inferred from the present data (nor methodology), we cannot state whether or not these differences are (partly) of genetic or (entirely) of environmental origin. That is, we cannot state whether genetic theories or culture-only theories are correct.

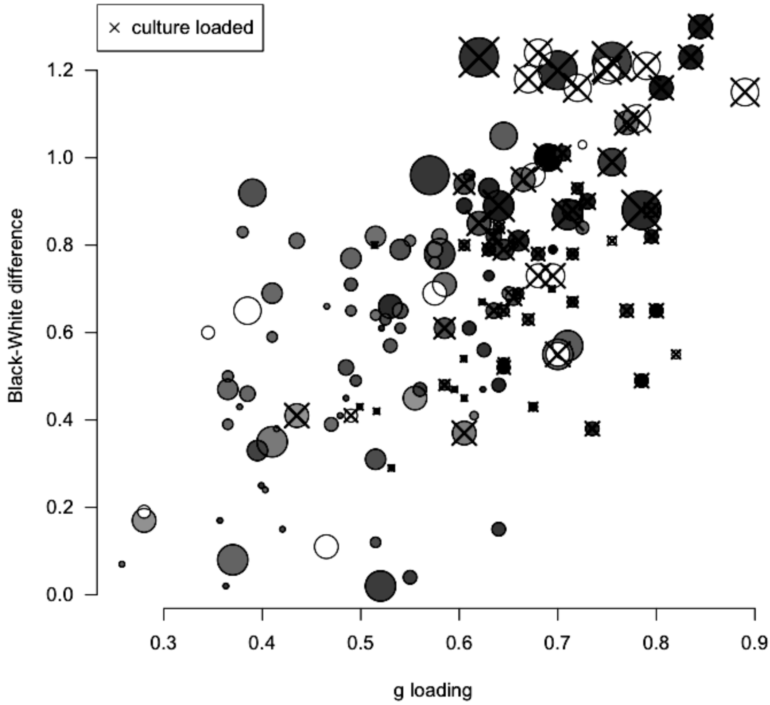

Figure 4.5 Scatterplot of Black-White Differences set out against loadings on the first principal factor for each study apart. The circles denote the data points; their size (area) represents the subject sample size; the darkness of the shade is proportional to the tests' reliabilities (circles are rendered in white if the reliability of the test is unknown).

Figure 4.6 Black-White differences are the largest on culture loaded subtests (left pane). Those tests are the most *g* loaded (middle pane) and the most reliable (right pane).

4.5 Discussion

We showed that Rushton & Jensen (2010a, p.214)'s hypothesis that "if population group differences are greater on the more *g* loaded and more heritable subtests, it implies they have a genetic origin" is invalid. Specifically we agree that a particular genetic origin of black-white differences would imply these correlations. However, we cannot infer a genetic origin from the correlations. This is because they are necessary also consistent with an environmental origin of black-white differences (hence with culture-only theory). The method on which Rushton & Jensen based their inferences is non-informative as applied in their editorial. The correlations have no necessary bearing on the origin of group differences.

Additionally, we reexamined Black-White data (Jensen, 1985; 1987) in the light of the theory that they favored, i.e., biological *g* theory (as we have denoted it). We provided evidence for an association that had not been discussed or clarified by the authors, namely the positive correlation between the subtests' Black-White mean differences and their cultural loading. Culture loaded test (language, knowledge, and achievement or scholastic tests) showed large *g* loadings and mean IQ differences, larger than culture reduced tests. The more cultural aspects of a test differentiate between people, the larger were the Black-White mean differences. This was inconsistent with biological *g* theory, because this theory predicted the opposite. Even if Black-White differences are indeed to some extend due to genetic differences, biological *g* theory does not explain why they are more pronounced on culture loaded tests than culture reduced tests.

Black-White differences were thus smaller on tests that measure cognitive processing than on tests that measure achievement ('on which races can grow apart as a result of dissimilar experiences', Rushton, 1998, p. 222). Achievement is certainly not fixed, but the result of past cognitive functioning and environmentally influenced learning experiences that take place throughout the course of development. In explaining the Black-White data adequately a developmental perspective is required; the use of cross-sectional data is insufficient. We advocate longitudinal designs and latent growth modeling, incorporating measured variables that replace in part or entirely the latent variables.

We entertain the possibility that *g* loadings are a function of environmental (culturally influenced) demands (Dickens & Flynn, 2001; Dickens 2008). Dickens simulated data using a model in which the more societal valued a (heritable) cognitive skill was, the more that skill was trained. He showed that, in this model, the most societal valued skills became the most *g* loaded and most heritable. In these simulations, interindividual differences were highly heritable, while intraindividual differences (growth) was large. The simulations provide an interesting perspective on individual and group differences. It implies that any initial environmental or genetic differences, large or small, can increase to large individual and group differences. This increase is due to a dynamical interplay between individuals' heritable characteristics and their environment. Such interplay will lead to a gene-environment correlation. We hypothesize that gene-environment correlation is higher for culture loaded ('crystallized') tests than culture reduced tests. Because gene-environment correlation can be incorporated in behavior genetic models (Purcell, 2002), this hypothesis is amenable to empirical investigation.

We conclude that the actual contributions of genes and environment to group differences cannot be determined or inferred from the method of correlated vectors. To date, whether and how genetic and environmental influences contribute to the Black-White differences remains unknown. This will remain unknown until the genetic and environmental variables that influence intelligence have been identified and the Black-White differences are regressed on them - preferably in a causal sense.

Table 4.1 Subtests' characteristics used in the statistical analyses

			Characteristic					
Test Number	Battery	Test name	Cultural load	*g* loading	Black-White Difference	Reliability	*g* loading*	Black-White Difference*
1	WISC-R	Information	1	.652	0.671	.778	.740	0.760
2	WISC-R	Similarities	1	.645	0.769	.810	.716	0.855
3	WISC-R	Arithmetic	1	.609	0.609	.770	.694	0.694
4	WISC-R	Vocabulary	1	.681	0.840	.826	.749	0.924
5	WISC-R	Comprehension	1	.547	0.648	.702	.653	0.773
6	WISC-R	Digit Span	0	.507	0.394	.712	.601	0.467
7	WISC-R	Picture Completion	0	.531	0.697	.770	.605	0.795
8	WISC-R	Picture Arrangement	0	.545	0.732	.726	.640	0.859
9	WISC-R	Block Design	0	.591	0.806	.846	.642	0.876
10	WISC-R	Object Assembly	0	.531	0.792	.700	.635	0.946
11	WISC-R	Coding	0	.334	0.343	.675	.407	0.417
12	WISC-R	Mazes	0	.404	0.729	.720	.476	0.860
13	WISC-R	Tapping Span	0	.387	0.330	.800	.432	0.369
14	CGP	Vocabulary	1	.687	0.913	.900	.724	0.962
15	CGP	Picture Number	0	.378	0.641	-	-	-
16	CGP	Reading	1	.759	0.928	.810	.844	1.031
17	CGP	Letter Groups	0	.644	0.985	.739	.749	1.145
18	CGP	Math	1	.785	1.090	-	-	-
19	CGP	Mosaic Comparisons	0	.398	0.882	.770	.453	1.005
20	SAT	Verbal	1	.886	1.150	-	-	-
21	SAT	Math	1	.792	1.210	-	-	-
22	ACT	English	1	.722	1.160	-	-	-
23	ACT	Social Studies	1	.752	1.200	-	-	-
24	ACT	Science Reading Comprehension	1	.693	1.240	-	-	-
25	ACT	Math	1	.685	1.180	-	-	-
26	Benttler Gestalt	Form Perception	0	.572	0.690	-	-	-
27	ITPA	Auditory-Vocal Association	0	.678	0.960	-	-	-
28	-	Draw a Man	0	.460	0.110	-	-	-
29	WRAT	Spelling	1	.694	0.730	-	-	-
30	WRAT	Reading	1	.678	0.730	-	-	-
31	WRAT	Arithmetic	1	.702	0.550	-	-	-
32	ASVAB	General Science	1	.836	1.230	.860	.901	1.326
33	ASVAB	Arithmetic Reasoning	1	.813	1.160	.870	.872	1.244
34	ASVAB	Vocabulary	1	.840	1.300	.860	.906	1.402
35	ASVAB	Paragraph Comprehension	1	.763	1.080	.680	.925	1.310
36	ASVAB	Numerical Operations	1	.657	0.950	.710	.779	1.060
37	ASVAB	Coding Speed	0	.559	0.960	.820	.617	1.350
38	ASVAB	Auto-Shop Information	1	.615	1.230	.830	.675	0.960
39	ASVAB	Mathematics Knowledge	1	.788	0.880	.840	.859	0.949

Table 4.1 (Continued)

			Characteristic					
Test Number	Battery	Test name	Cultural load	*g* loading	Black-White Difference	Reliability	*g* loading*	Black-White Difference*
40	ASVAB	Mechanical Comprehension	1	.707	1.200	.830	.776	0.867
41	ASVAB	Electronics Information	1	.758	1.220	.800	.847	1.364
42	GATB	Vocabulary	1	.640	0.890	.860	.690	0.960
43	GATB	Numerical	1	.710	0.870	.840	.775	1.317
44	GATB	Spatial	0	.580	0.780	.810	.644	0.867
45	GATB	Form Perception	0	.700	0.550	.730	.819	0.644
46	GATB	Name Comparison	0	.710	0.570	.750	.820	0.658
47	GATB	Motor Coordination	0	.520	0.020	.810	.578	0.022
48	GATB	Finger Dexterity	0	.410	0.350	.670	.501	0.428
49	GATB	Manual Dexterity	0	.370	0.080	.730	.433	0.094
50	K-ABC	Hands Movements	0	.512	0.549	.760	.587	0.630
51	K-ABC	Number Recall	0	.528	0.037	.810	.587	0.041
52	K-ABC	Word Order	0	.594	0.153	.820	.656	0.169
53	K-ABC	Gestalt Closure	0	.455	0.367	.710	.540	0.436
54	K-ABC	Triangles	0	.606	0.610	.840	.661	0.666
55	K-ABC	Matrix Analogies	0	.642	0.478	.850	.697	0.519
56	K-ABC	Spatial Memory	0	.552	0.467	.800	.617	0.522
57	K-ABC	Photo Series	0	.593	0.513	.820	.655	0.567
58	K-ABC	Faces and Places	1	.704	0.366	.840	.768	0.400
59	K-ABC	Arithmetic	1	.755	0.817	.870	.810	0.876
60	K-ABC	Riddles	1	.781	0.853	.860	.842	0.920
61	K-ABC	Reading (Decoding)	1	.747	0.479	.920	.779	0.500
62	K-ABC	Reading (Comprehension)	1	.772	0.620	.910	.810	0.650
63	WAIS	Digit Span	0	.627	0.410	.680	.760	0.497
64	Lorge-Thorndike	Sentence Completion	1	.818	0.550	-	-	-
65	Raven	Matrices	0	.731	1.030	-	-	-
66	Ammons	Picture Vocabulary	1	.750	0.810	-	-	-
67	WAIS	Information	1	.678	0.430	.910	.710	0.451
68	WAIS	Coding	0	.698	0.790	.920	.727	0.824
69	CGP	Sentences	1	.680	0.780	.840	.742	0.851
70	CGP	Integrative Reasoning	0	.722	0.840	.730	.845	0.983
71	CGP	Intersections	0	.285	0.190	-	-	-
72	CGP	Information About Technology	1	.479	0.410	-	-	-
73	CGP	Algebra	1	.628	0.790	.880	.670	0.670

Table 4.2

Correlation	Cultural load	*g* loading (N = 73)	Black-White Differences (N = 73)
Cultural load	1.00	0.67	0.53
g loading (N=54)	0.63	1.00	0.58
Black-White Differences (N=54)	0.49	0.57	1.00

Note: *g* loadings and Black-White standardized mean differences before (upper triangle) and after (lower triangle) correction for attenuation. N denotes the number of subtests.