Rushton, J.P. (2003). Race differences in g and the "Jensen Effect". In H. Nybörg (Ed.), The scientific study of general intelligence: A tribute to Arthur R. Jensen. (pp. 147-186). London: Elsevier.

Chapter 9

Race Differences in *g* and the "Jensen Effect"

J. Philippe Rushton

1. The Spearman-Jensen Hypothesis

Jensen (1980: 535) formally designated the view that Black-White differences were largely a matter of *g* as "Spearman's hypothesis", because Spearman (1927: 379) was the first to suggest it. Subsequently, Osborne (1980a) dubbed it the "Spearman-Jensen hypothesis" because it was Jensen who brought Spearman's hypothesis to widespread attention, and it was Jensen who did all the empirical work confirming it. More recently, to honor one of the great psychologists of our time, Rushton (1998) proposed that the term "Jensen Effect" be used whenever a significant correlation occurs between *g*-factor loadings and any variable, X; otherwise there is no name for it, only a long explanation of how the effect was calculated. Jensen Effects are not omnipresent and their absence can be as informative as their presence. For example, the "Flynn Effect" (the secular rise in IQ) is probably not a Jensen Effect because it does not appear to be on *g*.

The Black-White difference on the *g*-factor is the best known of all the Jensen Effects. The reason Jensen pursued Spearman's (1927) hypothesis was because it so exquisitely solved a problem that had long perplexed him. The average 15 to 18 IQ point difference between Blacks and Whites in the U.S. had not changed since IQ testing began almost 100 years ago. But Jensen (1969a) noted that the race differences were markedly smaller on tests of rote learning and short-term memory than they were on tests of abstract reasoning and transforming information. Moreover, culture-fair tests tended to give Blacks slightly *lower* scores than did conventional tests, as typically did non-verbal tests compared with verbal tests. Furthermore, contrary to purely cultural explanations, race differences could be observed as early as three years of age, and controlling for socioeconomic level only reduced the race differences by 4 IQ points.

Jensen (1968) initially formalized these observations in his so-called Level I-Level II theory. Level I tasks were those that required little or no mental manipulation of the input in order to arrive at the correct response whereas Level II tasks required mental manipulation. A classic example of Level I ability is Forward Digit Span in which people recall a series of digits in the same order as that in which they are presented. A

The Scientific Study of General Intelligence: Tribute to Arthur R. Jensen

ISBN: 0-08-043793-1

clear example of a Level II task is Backward Digit Span in which people recall a series of digits in the *reverse* order to that in which they were presented. Jensen found that Black-White differences were twice as large for Backward as for Forward Digit Span.

After Jensen (1980) re-read Spearman, he realized that the Black-White differences (and his Level I-Level II formulation) were specific examples of the more general hypothesis proposed by Spearman (1927: 379), namely that the Black-White difference "was most marked in just those [tests] which are known to be saturated with *g*". It was Spearman (1904, 1927), of course, who had generated the seminal concept of *g* in the first place. The *g* factor, derived from factor analysis of the correlations among a number of tests of mental abilities, is typically the largest factor.

To test Spearman's hypothesis, Jensen developed the *method of correlated vectors*. Essentially, this method correlates the standardized Black-White mean differences on a set of cognitive tests (a vector of scores, i.e. possessing both direction and quantity), with the tests' *g* loadings (a second vector of scores). A positive and substantial correlation provides support for Spearman's hypothesis. The rationale is straightforward: if *g* is the main source of between- *and* within-group differences, then there should be a positive relationship between a test's *g*-loading and the Black-White difference on the test; the more *g*-loaded the test, the greater the Black-White difference on that test. A methodological corollary is the prediction that when the point-biserial correlations of race (Black-White) with a number of diverse cognitive tests are entered into the total matrix of correlations among all the tests, the race variable will have its largest loading on the general factor of the correlation matrix.

According to Jensen (1998: 372–373), an ideal test of Spearman's hypothesis using the method of correlated vectors, must meet several methodological requirements. These are: (1) the samples being compared must be representative of their respective populations; (2) the samples being compared must be large enough to overcome the sampling error of the correlations among tests; (3) the samples being compared must not be selected on the basis of any *g*-loaded criterion; (4) the *g* factor should be extracted from enough tests to be reliable, as would be indicated by high coefficients of congruence in independent samples from the same population; (5) any test showing psychometric test bias for the groups being compared must be excluded; (6) the tests must be sufficiently diverse to allow significant differences between their *g* loadings; (7) the scores must be corrected for reliability; (8) the *g* values must be computed separately in the different groups; (9) the scores must measure the same latent traits in the different groups (i.e. the vector of *g* loadings extracted separately from each group must show a high congruence coefficient); and (10) the hypothesis must be tested for statistical significance by both Pearson's *r* and Spearman's rank-order correlation, rho.

As also noted by Jensen, tests of Spearman's hypothesis are necessarily stringent because the degrees of freedom used for statistical rejection of the null hypothesis are based on the number of pairs of variables in the correlated vectors (e.g. 13 sub-tests from the Wechsler Scales) and not on the subject sample size. It is also worth emphasizing that Spearman's hypothesis concerns the *relative* magnitude of the group difference across various tests that differ in their *g* loadings and not the *absolute* magnitude of group differences. It is therefore conceptually independent of any secular trend in absolute test scores, viz., the Flynn Effect (discussed below).

Jensen summarized his early tests of Spearman's hypothesis and responded to the open-peer commentary in *Behavioral and Brain Sciences* (Jensen 1985, 1987). Chapter 11 of *The g Factor* (1998) describes his subsequent analysis of 17 independent data sets of nearly 45,000 Blacks and 245,000 Whites derived from 171 psychometric tests (see Figure 9.1). The *g* loadings consistently predicted the relative magnitude of the Black-White differences ($r = 0.63$; Spearman rho $= 0.71$, $P < 0.05$) on the various tests. Spearman's hypothesis was borne out even among 3-year-olds administered eight subtests of the Stanford-Binet, where the rank correlation between *g* loadings and the Black-White differences was 0.71 ($P < 0.05$; Peoples *et al.* 1995).

These *g* related race differences are not due to factors such as the reliability of the test, social class differences, or the inevitable result of factor analysis. Indeed, the Spearman-Jensen hypothesis applies even to the *g* factor extracted from reaction time (RT) performance on elementary cognitive tasks. For example, in the "odd-man-out" task (Frearson & Eysenck 1986), 9- to 12-year-olds are asked to decide which of several lights is illuminated and then move their hand to press a button to turn that light off. All children can perform this and other tasks in less than 1 second, but children with higher IQ scores perform faster than do those with lower scores, and White children, on average, perform faster than do Black children (Vernon & Jensen 1984; Jensen 1993). The correlations between the *g* loadings of these types of RT tasks and the Black-White differences range from 0.70 to 0.81.

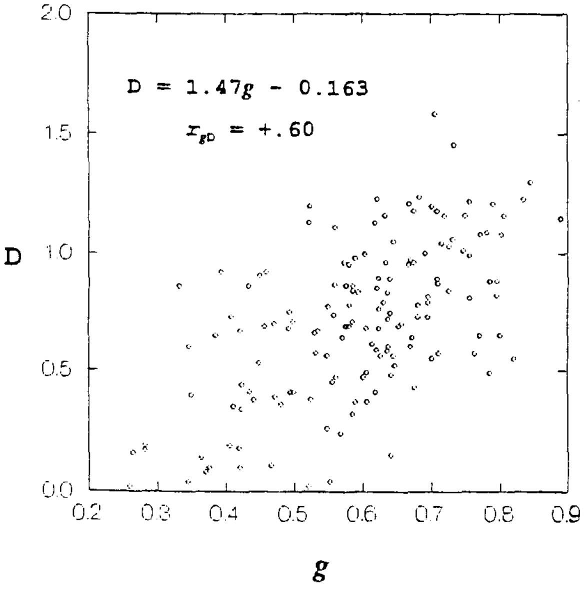

Figure 9.1: Scatter diagram of the correlation between the *g* loadings and the standardized mean White-Black differences (*D*) on 149 psychometric tests. With the tests' reliability coefficients partialled out, the correlation is 0.63 ($P < 0.001$). (After Jensen 1998: 378).

When Jensen examined East Asian-White comparisons using these same reaction time measures, the direction of the correlation was *opposite* to that in the Black-White studies, with East Asians scoring higher in *g* than Whites (Jensen & Whang 1993, 1994). Dozens of other studies indicate an East Asian advantage on conventional IQ tests (Lynn 1991, 1997; Lynn & Vanhanen 2002; see also Chapter 8 in this volume). So far, however, only one study seems to have looked at East Asian-White differences on conventional psychometric tests as a direct function of their *g* loadings. Nagoshi *et al.* (1984) administered 15 cognitive tests to two generations of Americans of Japanese, Chinese and European ancestry. Of the four reported correlations between *g* loadings and ethnic group differences, only one was significant, albeit in the predicted direction. It is interesting to note, in light of the above, that in an early reply to a charge of "white supremacy", Jensen wrote (1969b: 240) "[I]f I were asked to hypothesize about race differences in what we call *g* or abstract reasoning ability, I should be inclined to rate Caucasians on the whole somewhat below Orientals, at least those in the United States."

2. Studies of Race Differences

2.1. U.S. Black-White Differences in g Since 1998

Following publication of *The g Factor* (Jensen 1998), Nyborg & Jensen (2000) provided further evidence for the generality of the Spearman-Jensen hypothesis. They analyzed a unique battery of 19 highly diverse cognitive tests made available by the Centers for Disease Control (CDC: 1988) from an archival data set of 4,462 males who had served in the U.S. Armed Forces. Since approximately half had served in the Vietnam War, the CDC made the data available for researchers to examine the effects of exposure to toxic substances such as Agent Orange used during that war. The sample was fairly representative of the U.S. population in terms of race, education, income and occupation, though the lower tail of the ability distribution was truncated because those who scored below the 10th percentile on a pre-induction general aptitude test were excluded from military service.

Five of the tests were administered at the time of induction into the armed forces while the others were given approximately 17 years later. The battery included the Grooved Pegboard Test, the Paced Auditory Serial Addition Test, the Rey-Osterrieth Complex Figure Drawing Test, the Wisconsin Card Sorting Test, the Wide Range Achievement Test, the California Verbal Learning Test, the General Information Test, and various sub-tests from the Wechsler Adult Intelligence Scale-Revised, and the Army Classification Battery. The *g* factor was extracted using different methods. Spearman's hypothesis was confirmed regardless of procedure, with an average correlation between the race differences on a test and its *g* loading across the extraction procedures of 0.81. This value is higher than the average correlation of 0.62 for all previous studies. Nyborg & Jensen (2000) concluded that Spearman's original conjecture about the Black-White difference on the *g* factor "should no longer be regarded as just an hypothesis but as an empirically established fact" (p. 599).

Subsequently, Roth *et al.* (2001) carried out a meta-analytic review that extended the range of the 1.1 standard deviation effect size of Black-White IQ differences to college and university application tests such as the Scholastic Aptitude Test (SAT; $N = 2.4$ million) and the Graduate Record Examination (GRE; $N = 2.3$ million), as well as to tests for job applicants in corporate settings ($N = 0.5$ million), and in the military ($N = 0.4$ million). Their review also noted that in any particular work settings where the Black-White difference was less than 1.1 standard deviations, it resulted from selection factors and restriction of range (as in studies of complex versus simple jobs, within-versus across-jobs, and of job incumbents versus applicants). Roth *et al.* (2001) report that the tests with the greatest *g* loadings showed the largest Black-White differences, again confirming Spearman's hypothesis. Since IQ scores are the best predictor of economic success in Western society (Schmidt & Hunter 1998), group differences in IQ scores help to explain the differences in societal outcomes (Gottfredson 1997).

Jensen's (in press) most recent study of Black-White differences compared 8- to 13-year-old pupils (877 White, 855 Black) on up to 17 tests. As usual, race differences at all ages were mainly on the *g* factor. Age differences also showed a Jensen Effect, that is, older children had higher levels of *g*, rather than just more acquired knowledge. The race effect (which was equal to about two years of chronological age) differed from the age effect by being greater on tests with higher *g* loadings (such as verbal and figural reasoning) whereas the age effect was greater on tests with lower *g* loadings (such as digit span memory). This age x race interaction suggests that for both racial groups, cognitive development during childhood involves mental growth factors other than *g*, while the Black-White difference at any given age is almost exclusively a matter of *g*.

2.2. Studies of g in South Africa

Since the studies so far described were all carried out in the U.S., it might be argued that the results are solely due to local conditions. However, the race differences on the *g* factor have been confirmed in several independent studies carried out in sub-Saharan Africa (this section) and in the Netherlands (next section). It is surprising that only recently has research on the *g* factor been carried out in Africa because the low mean test scores obtained there are of considerable interest. They came to widespread attention in the U.S. with Richard Herrnstein and Charles Murray's best-selling 1994 book *The Bell Curve* (and my own 1995 *Race, Evolution, and Behavior*). *The Bell Curve* considered the hypothesis that: "The African black population has not been subjected to the historical legacy of American black slavery and discrimination and might therefore have higher scores". However, after examination of the literature, the IQ for Black Africans turned out to be, on average, substantially below African Americans, who are of mixed Black and White ancestry.

Both *The Bell Curve* and *Race, Evolution, and Behavior* cited a 1991 review by Richard Lynn of 11 studies from West, East, Central, and Southern Africa which reported an average IQ of 70 (median = 75), 15 points (1 standard deviation (SD)) lower than the mean of 85 typically found for Black Americans and 30 points (2 SDs) lower than the mean of 100 typically found for Whites. Subsequent reviews by Lynn (1997;

see also his Chapter 8 in this volume) that examined additional published studies have confirmed a mean African IQ of 70. In their recent book *IQ and the Wealth of Nations*, Lynn & Vanhanen (2002) reviewed over two-dozen studies from the African sub-continent finding the lowest recorded IQ scores in the world.

University students in South Africa also achieve low mean test scores. Sixty-three undergraduates at the all-Black universities of Fort Hare, Zululand, the North, and the Medical University of South Africa were given the Wechsler Adult Intelligence Scale-Revised (WAIS-R) and found to have a full scale IQ of 77 (Avenant 1988; cited by Nell 2000: 26–28). A study at the University of Venda in South Africa's Northern Province by Grieve & Viljoen (2000) found 30 students in 4th-year law and commerce averaged a score of 37 out of 60 on the Standard Progressive Matrices. This is equivalent to an IQ of 78 on the 1993 U.S. norms (Raven *et al.* 1990: 98; 1998: 77). A study at South Africa's University of the North by Zaaiman, van der Flier & Thijs (2001) found the highest scoring African sample to that date — 147 first-year mathematics and science students who scored 52 out of 60 on the Standard Progressive Matrices, with an IQ equivalent of 100. Their relatively high mean score may have been because they were mathematics and science students, and also because they had been specially selected for admission to the university from a pool of 700 applicants on the basis of a mathematics and science selection test.

Lynn & Owen (1994) were the first to explicitly test Spearman's hypothesis in sub-Saharan Africa. They administered the Junior Aptitude Test, a paper-and-pencil test, consisting of ten sub-tests (four verbal, six nonverbal), to 1,056 White, 1,063 Indian, and 1,093 Black 16-year-old high school students in South Africa. They found a 2 SDs difference between the Africans and Whites (yielding an average African IQ of about 70) and a 1 SD difference between the Whites and Indians (yielding an average Indian IQ of 85). Lynn and Owen tested Spearman's hypothesis and found the African-White differences correlated 0.62 ($P < 0.05$) with the g-factor extracted from the African sample (although only 0.23 with g extracted from the White sample). However, unlike the African-White differences, the White-Indian differences they found were *not* on the g factor.

Six studies by Rushton, Skuy, and other colleagues in Southern Africa also support Spearman's hypothesis, including of university students. In the first study, Rushton & Skuy (2000) gave 309 17- to 23-year-old first-year psychology students at the University of the Witwatersrand in Johannesburg the untimed Standard Progressive Matrices. The 173 African students solved an average of 44 of the 60 problems, while the 136 White students solved an average of 54 of the 60 problems. These scores placed the African students at the 14th percentile and the White students at the 61st percentile, which yielded IQ equivalents of 84 and 104, respectively. Because the total score on the Standard Progressive Matrices is an excellent measure of g, Rushton and Skuy used the item-total correlations as an estimate of each item's g loading and found that item g loadings showed a significant positive correlation with the standardized differences in the percentage of Africans and Whites passing the same items. These Jensen Effects were found using both the African item-total correlations, $r = 0.39$ ($P < 0.01$, $N = 58$, with rho $= 0.43$, $P < 0.01$), and the White item-total correlations, $r = 0.34$ ($P < 0.01$, $N = 46$, rho $= 0.41$, $P < 0.01$).

The second study (Rushton 2001) analyzed ten sub-tests of the Weschler Intelligence Scale for Children-Revised (WISC-R) from data published by Skuy *et al.* (2001) on 154 Black South African high school students from Johannesburg. Table 9.1 presents the African means and SDs for the various WISC-R sub-tests. The table shows the African mean scores are 1 to 2 SDs below American norms. The mean score for Whites was set at the U.S. standardization sample mean of 10 (which included African Americans). The mean African-White differences were then calculated and also expressed in SD units, using the African SDs. When the *g* loadings from the WISC-R national standardization data were extracted they correlated $r = 0.77$ ($P < 0.05$) with the standardized African-White differences, thereby showing the Jensen Effect. For many of the African students, English was not their first language. However, the Jensen Effect remained even after the Vocabulary sub-test was excluded from the data (in Table 9.1), and the mean of the 11 other sub-tests substituted in its place ($r = 0.66$, $P < 0.05$). Nor did the Jensen Effect disappear if *g* was extracted from the African rather than from the White standardization sample ($r = 0.60$, $P < 0.05$), or if Spearman's rho was used instead of Pearson's *r* to measure the magnitude of the correlation (rhos = 0.74, 0.74, respectively, $P\text{s} < 0.005$).

The third study (Rushton 2002) re-analyzed published data from Owen (1992) who had given the Standard Progressive Matrices in South Africa without time limits to 1,056 White, 1,063 Indian, 778 Colored, and 1,093 Black 14-year-olds. Out of a total of 60 items, Owen (1992) found the Whites averaged 45 correct, East Indians, 42, Coloreds, 37, and Blacks 28, placing them at the 57th, 42nd, 19th, and 5th percentiles, yielding IQ equivalents of 103, 97, 87 and 75 on the 1993 U.S. norms. Importantly, Owen found that the item-total test score correlations predicted the pass rate differences between the ethnic groups on these same items and concluded that this indicated an absence of test bias. Rushton proposed a stronger inference, that all the group differences (viz., White-African, White-Colored, White-Indian, Indian-African, Indian-Colored, Colored-African) were primarily on *g*. To test this possibility, he carried out a purely non-parametric re-analysis of Owen's data and found that, indeed, the more highly correlated an item was with *g* (the item-total correlation), the more it predicted the differences among the (now ranked) item pass rates for Whites, Indians, Coloreds, and Africans, (Spearman's rhos from 0.35 to 0.85; all $P\text{s} < 0.01$). The effects remained regardless of the ethnic group from which the item *g*-loadings were taken.

In the fourth study, I teamed up with Arthur Jensen himself (Rushton & Jensen 2003) to analyze a set of data published by Zindi (1994), an African Zimbabwean educational psychologist. Zindi had reported data on 10 sub-tests of the WISC-R for 204 Black Zimbabwean children with a total IQ score of 70, a difference of nearly 2 SDs below White norms. Because the sub-test correlations were not available, Rushton and Jensen compared the Zimbabwean means and SDs against those for 1,868 White Americans from the U.S. standardization sample. A principal factor analysis of the correlation matrix was carried out for the White standardization sample, along with the point-biserial correlation of the African-White standardized differences on each sub-test, a measure of the racial "effect size". Table 9.2 shows the loadings of the African-White variable and of the 10 WISC-R sub-tests on the *g* factor, and on the next three largest unrotated principal factors (regardless of sign and whether the eigenvalues were less than 1). The *g* loading is considerably larger than the largest non-*g* factor, and the ratio

Table 9.1: Means and standard deviations (*SD*s) of African secondary school students aged 13 to 15 years on sub-tests of the WISC-R and U.S.-African differences (After Skuy *et al.* 2001; Rushton 2001).

WISC-R Scale	**African mean**	**African *SD***	**U.S.-African difference**	***z*-score difference**	***g* for U.S.**	***g* for Africa**	**Reliability**
Information	4.66	2.33	5.34	2.29	0.67	0.65	0.85
Picture Completion	7.06	2.28	2.94	1.29	0.51	0.57	0.77
Similarities	4.89	2.32	5.11	2.20	0.67	0.62	0.81
Picture Arrangement	6.42	2.68	3.58	1.34	0.49	0.49	0.73
Arithmetic	6.01	2.20	3.99	1.81	0.57	0.60	0.77
Blocks	6.58	2.49	3.42	1.37	0.65	0.61	0.85
Vocabulary	2.85	1.64	7.15	4.36	0.72	0.71	0.86
Object Assembly	6.29	2.81	3.71	1.32	0.50	0.53	0.70
Comprehension	4.79	2.46	5.21	2.12	0.60	0.61	0.77
Coding	6.18	2.25	3.82	1.70	0.37	0.36	0.72
Digits	6.93	2.64	3.07	1.16	0.44	0.59	0.78
Mazes	7.60	2.90	2.40	0.83	0.37	0.45	0.72

U.S. standardization sample mean = 10, *SD* = 3.

Table 9.2: Principal factor analysis of correlation matrix for 1,868 whites plus the African-White standardized effect size. (After Rushton & Jensen in press).

WISC-R Sub-tests		**Non-*g* factors**		
	g	**1>**	**2>**	**3>**
Information	0.724	–0.274	0.106	–0.164
Similarities	0.711	–0.217	0.144	0.002
Arithmetic	0.592	–0.138	–0.077	–0.195
Vocabulary	0.772	–0.313	0.032	0.002
Comprehension	0.705	–0.247	–0.084	0.319
Picture completion	0.556	0.158	0.194	0.066
Picture arrangement	0.517	0.078	0.157	–0.023
Block design	0.697	0.311	0.154	–0.112
Object assembly	0.582	0.426	0.188	0.062
Coding	0.429	0.107	–0.365	–0.204
Race	0.882	0.236	–0.389	0.124
Percent Total Variance Explained	44.073	6.164	4.131	2.228

of g variance to all non-g variance is 3.5 to 1. The race variable has a larger g loading than do any of the sub-test variables. It appears to reflect almost pure g. Fully 77% of the between-group race variance can be attributed to a single source, namely g.

The fifth study (Rushton *et al.* 2002) gave the Standard Progressive Matrices to an academically select population of 342 17- to 23-year-old first-year engineering students (198 Africans, 58 Indians, 86 Whites) in the Faculties of Engineering and the Built Environment at the University of the Witwatersrand. Out of the 60 problems, the African students solved an average of 50, the Indian students, 53, and the White students, 56, placing the Africans at the 41st percentile, the Indians at the 50th, and the Whites at the 75th, with IQ equivalents of 97, 105, and 110, respectively. Several analyses showed that even for this very select group, the standardized African-Indian-White differences were most pronounced on those items with the highest item-total correlations, indicating a difference in g. Indeed, the g-loadings showed cross-cultural generality; for example, item-total correlations calculated on the Indian students predicted the magnitude of the African-White differences. When the 60 items were aggregated into 10 "sub-tests" of six items each, the magnitude of the Jensen Effect was similar to that from studies based on whole sub-tests (median rho = 0.53).

In a sixth study, Rushton *et al.* (2003) gave the Advanced version of the Progressive Matrices to 294 of the same engineering students (187 Africans, 40 Indians, 67 Whites) as in the previous study. Out of the 36 problems, the African students solved an average of 22, the Indian students, 24, and the White students, 29, placing the Africans at the 57th percentile, the Indians at the 64th, and the Whites at the 86th, with IQ equivalents of 103, 106, and 117, respectively, making this the now highest scoring African sample on record. External validities were established, with both the African and the non-African students' scores on the Advanced Progressive Matrices predicting their scores on the Standard Progressive Matrices taken three months earlier (mean $r = 0.60$; mean $Ps < 0.01$) and their final examination marks taken three months later (mean $rs = 0.30$; mean $Ps < 0.01$). Once again, the standardized African-Indian-White differences were Jensen Effects, being most pronounced on those items with the highest item-total correlations. Moreover, the g-loadings again showed cross-cultural generality, with those calculated on the Indian students predicting the magnitude of the African-White differences.

2.3. *Ethnic Differences in g from The Netherlands*

Several studies of g-factor differences among various populations have been carried out in The Netherlands by Jan te Nijenhuis and his colleagues who compared the majority Dutch population with immigrants from the Third World who now comprise 6% of the Dutch population. About 40% of these immigrants came from the West Indies — the Netherland Antilles and Surinam. The Antilleans are predominantly of mixed African descent and the Surinamese are a diverse population of Creoles (mixed African, White, and Native American), East Indians, and individuals of Indonesian and Chinese descent. They speak Dutch as their first language. About 60% of the immigrants came from Turkey and Morocco. These are Caucasian, and do not have Dutch as their first language. The IQ scores of all immigrants averaged over 1 SD lower than did those of the Dutch majority, with the North Africans and Turks scoring lower than the Surinamese and Antilleans, especially on tests with a verbal component. Immigrant children tended to perform poorly in school and the adult unemployment rate was 20% for immigrants versus 7% for the total population (te Nijenhuis & van der Flier 2001).

In one study, te Nijenhuis & van der Flier (1997) compared the test results of all first generation immigrant job applicants to the Dutch Railways between 1988 and 1992 with those of a random representative sample of all the majority group applicants over the same time period. The Dutch version of the General Aptitude Test Battery (GATB), consisting of eight speeded sub-tests, showed similar alpha coefficients and covariance matrices in all groups and, apart from a Dutch language proficiency factor, there was no evidence of test bias. The Dutch-Immigrant differences correlated highly with the g factor extracted separately for each of the five groups, before and after correcting for unreliability in the measures. (The after-correction correlations were: Dutch-Surinamese, $r = 0.76$; Dutch-Antilleans, $r = 0.78$; Dutch-North Africans, $r = 0.82$; and Dutch-Turks, $r = 0.64$.)

The Spearman-Jensen hypothesis was also tested on this sample using safety aptitude measures (ability to concentrate and sensori-motor coordination ability) that are important predictors for accident-related criteria for this sample of engine drivers, guards, train traffic controllers, bus drivers, shunters, and railway station assistants (te Nijenhuis 1997). The safety aptitude scores were consistently lower for all the immigrant groups than for the Dutch majority group. Group differences on the safety aptitude tests correlated $r = 0.81$ with their g loadings from the GATB, indicating that the group differences in safety aptitude were largely a function of g, i.e. they were Jensen Effects.

Subsequently, te Nijenhuis *et al.* (2000) gave the Dutch adaptation of the Differential Aptitude Test to 318 Dutch and 111 immigrant secondary school students (no details being available about the ethnic origin of the immigrant students). School grades and scholastic achievement test scores were used as criteria. On the g factor scores extracted from nine sub-tests, the mean of the immigrants was 1.14 SDs below that of the Dutch, with the pattern of g loadings similar for both. The immigrant group was also lower on the criteria measures, which were predicted equally well for both groups. Group differences in both test scores and educational achievement were predicted quite well by the g loadings of the various measures, making g the predominant factor accounting for the group differences.

In still another study, te Nijenhuis *et al.* (in press) examined the Revised Amsterdam Intelligence Test for Children (RAKIT), which consisted of 12 sub-tests. The study compared 604 Dutch children who constituted the nationally representative normative sample against 559 immigrant children who had been carefully selected to be generally representative of all immigrant children in the Netherlands. Little test bias was found; some differential prediction occurred, but its effects were small. The estimate of g as computed from the test showed strong predictive validity for most school subjects and standardized achievement tests. Moreover, the study confirmed Spearman's hypothesis that g is the predominant factor determining the size of the difference between the two groups.

In summarizing all the Dutch studies on the assessment of immigrants, te Nijenhuis & van der Flier (2001) stated that the lower scores of the immigrants could be generalized to the whole population of immigrants, yielding IQs under 100 for 84% of immigrants. However, second-generation immigrants were doing roughly one-third of an SD better than were first-generation immigrants, and the third generation may continue to improve their group's relative position. As many other West European countries also have immigrants from Third World countries, including former colonies, the Dutch findings may be generalizable.

2.4. The Flynn Effect May Not Be a Jensen Effect

Jensen Effects are not omnipresent and their absence can be as informative as their presence. For example, it is not universally true that all groups that differ, on average, in their overall score on a test battery will necessarily conform to the Spearman-Jensen hypothesis. A study in Spain by Colom *et al.* (2002), using the Spanish standardization sample of the Wechsler Adult Intelligence Scale (WAIS III), found that while 2 SDs divided the lowest IQ group (IQ = 84) from the highest IQ group (IQ = 112), Jensen Effects were not found on the 14 sub-tests: the people were apparently not less intelligent, merely less educated.

In a study by Lynn & Owen (1994) in South Africa, although there was a nearly 1 SD difference between Whites and East Indians, there was no correlation between *g* loadings and standardized mean differences on 10 sub-tests. Thus, it was not a Jensen Effect. (Several subsequent studies, however, found the Indian-White and Indian-African differences *were* on the *g* factor; Rushton 2002; Rushton *et al.* 2002). It is an interesting question which of the national differences documented in Lynn & Vanhanen's (2002) *IQ and the Wealth of Nations* are on the *g* factor.

Another apparent absence of the Jensen Effect is that shown for the secular increase in test scores. The Flynn Effect (sometimes also known as the Lynn-Flynn Effect) refers to the repeated demonstration by Flynn (1984, 1987, 1999; but see also Lynn 1982) that the average IQ in several countries has increased by about 3 points a decade over the last 50 years. The Flynn Effect seems to imply the 1 SD difference in the mean IQ between Blacks and Whites in the U.S. will simply disappear over time (Flynn 1999). However, analysis shows that the Flynn Effect is not on the *g* factor.

Table 9.3 (based on Rushton 1999) shows the results of a principal components analysis of the secular gains in IQ from the U.S., Germany, Austria and Scotland, along with Black-White IQ difference scores from the U.S., inbreeding depression scores from cousin marriages in Japan, and *g*-loadings from the standardization samples of the Wechsler Intelligence Scale for Children (WISC-R and WISC-III). The important findings are: (1) the IQ gains on the WISC-R and WISC-III form a cluster, showing that the secular trend is a reliable phenomenon; but (2) this cluster is *independent* of a second cluster formed by Black-White differences, inbreeding depression scores (a purely genetic effect), and *g*-factor loadings (a largely genetic effect). This analysis shows that the secular increase in IQ behaves in a different way than the mean Black-White IQ difference. The secular increase is unrelated to *g* and other heritable measures,

Table 9.3: Principal components analysis and varimax rotation for Pearson correlations of inbreeding depression scores, black-white differences, *g* loadings, and gains over time on the Wechsler intelligence scales for children with reliability partialled out. (After Rushton 1999).

Variables	Principal Components			
	Unrotated Loadings		Varimax Rotated Loadings	
	I	II	1	2
Inbreeding depression scores from Japan (WISC-R)	0.31	0.61	0.26	**0.63**
Black-White differences from the U.S. (WISC-R)	0.29	0.70	0.23	**0.72**
WISC-R *g* loadings from the U.S.	–0.33	0.90	–0.40	**0.87**
WISC-III *g* loadings from the U.S.	–0.61	0.64	–0.66	**0.59**
U.S. gains 1 (WISC to WISC-R)	0.73	–0.20	**0.75**	–0.13
U.S. gains 2 (WISC-R to WISC-III)	0.81	0.40	**0.77**	0.47
German gains (WISC to WISC-R)	0.91	0.03	**0.91**	0.11
Austria gains (WISC to WISC-R)	0.87	0.00	**0.86**	0.07
Scotland gains (WISC to WISC-R)	0.97	0.08	**0.96**	0.17
% of total variance explained	48.6	25.49	48.44	25.65

while the magnitude of the Black-White difference is related to heritable *g* and inbreeding depression.

Flynn's hypothesis that the "massive IQ gains over time" imply an environmental origin of race differences is not supported. Although the Flynn Effect does suggest that improving the environment, especially at the low end of the IQ distribution, can improve test scores, the cluster analysis show that effect is unrelated to the *g*-factor. Instead, *g* is associated with inbreeding depression, for which there is no non-genetic explanation, which implies strongly that *g* is less amenable to environmental manipulation.

Two Estonian studies confirm the finding that the Flynn Effect is not on the g factor (Must *et al.* in press (a), in press (b)). In the largest of these, Must *et al.* (in press (b)) analyzed ten sub-tests of the Estonian translation of the (U.S.) National Intelligence Test from comparable samples of 12- to 14-year-old children taken over a 60-year period (1936 to 1998). The loadings on the 1st Principal Component (which represents g) had a congruence coefficient across the 60-year time span of 0.996, but *negative* correlations of –0.24 to –0.54 with the 60-year Flynn Effect of sub-test gains (depending on how the sample was divided or how $+g$ was extracted).

In Spain, however, Colom *et al.* (2001), have reported a positive correlation ($r = 0.78$; $P < 0.05$) between g and the amount of generational change in two successive standardizations of the Spanish Differential Aptitude Test across 16-years. There were 10 samples of males and females for each of five sub-tests (Verbal Reasoning, Space Relations, Numerical Ability, Mechanical Reasoning, and Abstract Reasoning). But 5 of the 10 samples showed a generational *decrement* (their Table 1), so ambiguities in the study raise questions about its generality. Moreover, the magnitude of the Jensen Effect on the secular rise in IQ that Colom *et al.* found is relatively small compared to the Jensen Effect for race (about 0.50 SD compared to >1.00 SD for Black-White differences).

In summary, no one has yet found a Flynn Effect that approaches the magnitude of the Jensen Effect. All but the Colom *et al.* (2001) study (i.e. Rushton 1999; Must *et al.* in press (a), in press (b)) showed no Jensen Effect at all (or even negative correlations between secular gains and g loadings). The complete explanation for the secular rise in IQ remains one of the unsolved psychometric mysteries.

2.5. *Pushing Out the Envelope Even Further*

From the beginning, "Jensenism" did not stop within the U.S. or with IQ. For example, Jensen (1969a: 86) cited studies showing the early development of motor behavior in Black infants, with some Black samples at six months of age scoring nearly 1 SD above White norms. Paralleling the behavioral precocity, Jensen (1969a: 87) reported evidence of faster bone development in Black infants (established using X-rays) and earlier maturation of brain wave patterns (measured using EEGs). Soon after, Jensen (1973: 289–290) suggested that race differences in the production of two-egg twins, being most common among Blacks and least common among East Asians, with Caucasians intermediate, "may be a reflection of evolutionary age". In a long footnote, he wrote: "[T]he three racial groups lie on a developmental continuum on which the Caucasian group is more or less intermediate. A related fact is that there is an inverse relationship throughout the phylogenetic hierarchy between the tendency for multiple births and the prolongation of immaturity."

Many researchers were inspired by "Jensenism". Richard Lynn (1977, 1978, 1982) and Philip E. Vernon (1982) not only pushed the envelope, but extended the "outside of the envelope" and made the race-IQ debate international in scope with their findings that East Asians average higher on tests of mental ability than do Whites, whereas Caribbeans (and especially Africans) average lower. As Lynn's (1997, 2003;.Lynn &

Vanhanen 2002) most recent reviews show, East Asians, measured in North America and in Pacific Rim countries, typically average IQs in the range of 101 to 111; Whites in North America, Europe, and Australasia average IQs of 85 to 115, with an overall mean of 100; and Blacks, living south of the Sahara, in North America, in the Caribbean, and in Britain average IQs of 70 to 90.

As a budding sociobiologist, I too was inspired by Jensenism. It seemed to me that by its impact on diverse areas of behavioral science, Jensenism might help complete the Darwinian revolution. I began to review the international literature, studying not only IQ, but other behavioral traits like speed of physical maturation and longevity, personality and temperament, family structure and crime, and sexual behavior and fertility, and later brain size too (Rushton 1984a, 1984b, 1988). These studies culminated in a book *Race, Evolution, and Behavior* (Rushton 1995, 2000). East Asians are slower maturing, less fertile, less sexually active, with larger brains and higher IQ scores than Blacks, who tend towards the opposite in each of these areas. Whites fall between the other two groups (see Table 9.4). As Jensen (1984) elaborated in a commentary on my first (1984a) review, a network of such related evidence provides more opportunity for finding and testing alternative theories than does any single dimension drawn from the set.

Subsequently, I carried out experiments finding, for example, that the amount of inbreeding depression on 11 sub-tests of the WISC-R in Japan predicted the magnitude of the Black-White differences on the same sub-tests in the U.S. (Rushton 1989). Inbreeding depression, a purely genetic effect, was a sufficiently robust predictor to overcome generalization from the Japanese in Japan to Blacks and Whites in the U.S. There really is no explanation, other than a genetic one, for the correlation between inbreeding depression and Black-White differences.

I also examined the relation between intelligence and brain size, finding correlations of $r=0.20$ between IQ scores and simple head size measures and of $r=0.44$ with measures based on Magnetic Resonance Imaging (MRI) (Rushton & Ankney 1996). In one study, Rushton (1997) analyzed data from the enormous Collaborative Perinatal Project which recorded head circumference measurements and IQ scores from 50,000 children followed from birth to age seven (Broman *et al.* 1987). At birth, four months, one year, and seven years, the Asian American children averaged larger cranial volumes than did the White children, who averaged larger cranial volumes than did the Black children (Figure 9.2). Within each race, the children with the larger head sizes had the higher IQ scores. By age seven, the Asian American children averaged an IQ of 110, White children an IQ of 102, and Black children an IQ of 90. Since the Asian American children were the shortest in stature and the lightest in weight and the Black children were the tallest in stature and the heaviest in weight, these race differences in brain-size/IQ relations were not due to body size.

With adults, I used external head size measurements (length, width, height) to calculate cranial capacities from five large archival data sets. In the first of these studies, Rushton (1991) examined head size measures in 24 international military samples collated by the U.S. National Aeronautics and Space Administration. After adjusting for the effects of body height, weight, and surface area, it found the mean cranial capacity for East Asians was 1,460 and for Europeans 1,446 cm^3. The second (Rushton 1992)

Table 9.4: Relative ranking of races on diverse variables.

Variable	**East Asians**	**Whites**	**Blacks**
Brain size			
Autopsy data (cm^3 equivalents)	1,351	1,356	1,223
Endocranial volume (cm^3)	1,415	1,362	1,268
External head measures (cm^3)	1,356	1,329	1,294
Cortical neurons (billions)	13.767	13.665	13.185
Intelligence			
IQ test scores	106	100	85
Decision times	Faster	Intermediate	Slower
Cultural achievements	Higher	Higher	Lower
Maturation rate			
Gestation time	?	Intermediate	Earlier
Skeletal development	Later	Intermediate	Earlier
Motor development	Later	Intermediate	Earlier
Dental development	Later	Intermediate	Earlier
Age of first intercourse	Later	Intermediate	Earlier
Age of first pregnancy	Later	Intermediate	Earlier
Life-span	Longer	Intermediate	Shorter
Personality			
Activity	Lower	Intermediate	Higher
Aggressiveness	Lower	Intermediate	Higher
Cautiousness	Higher	Intermediate	Lower
Dominance	Lower	Intermediate	Higher
Impulsivity	Lower	Intermediate	Higher
Self-concept	Lower	Intermediate	Higher
Sociability	Lower	Intermediate	Higher
Social organization			
Marital stability	Higher	Intermediate	Lower
Law abidingness	Higher	Intermediate	Lower
Mental health	Higher	Intermediate	Lower
Administrative capacity	Higher	Higher	Lower
Reproductive effort			
Two-egg twinning (per 1000 births)	4	8	16
Hormone levels	Lower	Intermediate	Higher
Secondary sex characteristics	Smaller	Intermediate	Larger
Intercourse frequencies	Lower	Intermediate	Higher
Permissive attitudes	Lower	Intermediate	Higher
Sexually transmitted diseases	Lower	Intermediate	Higher

Note: From: Rushton, J. P. (1995). *Race, evolution, and behavior* (p. 5).

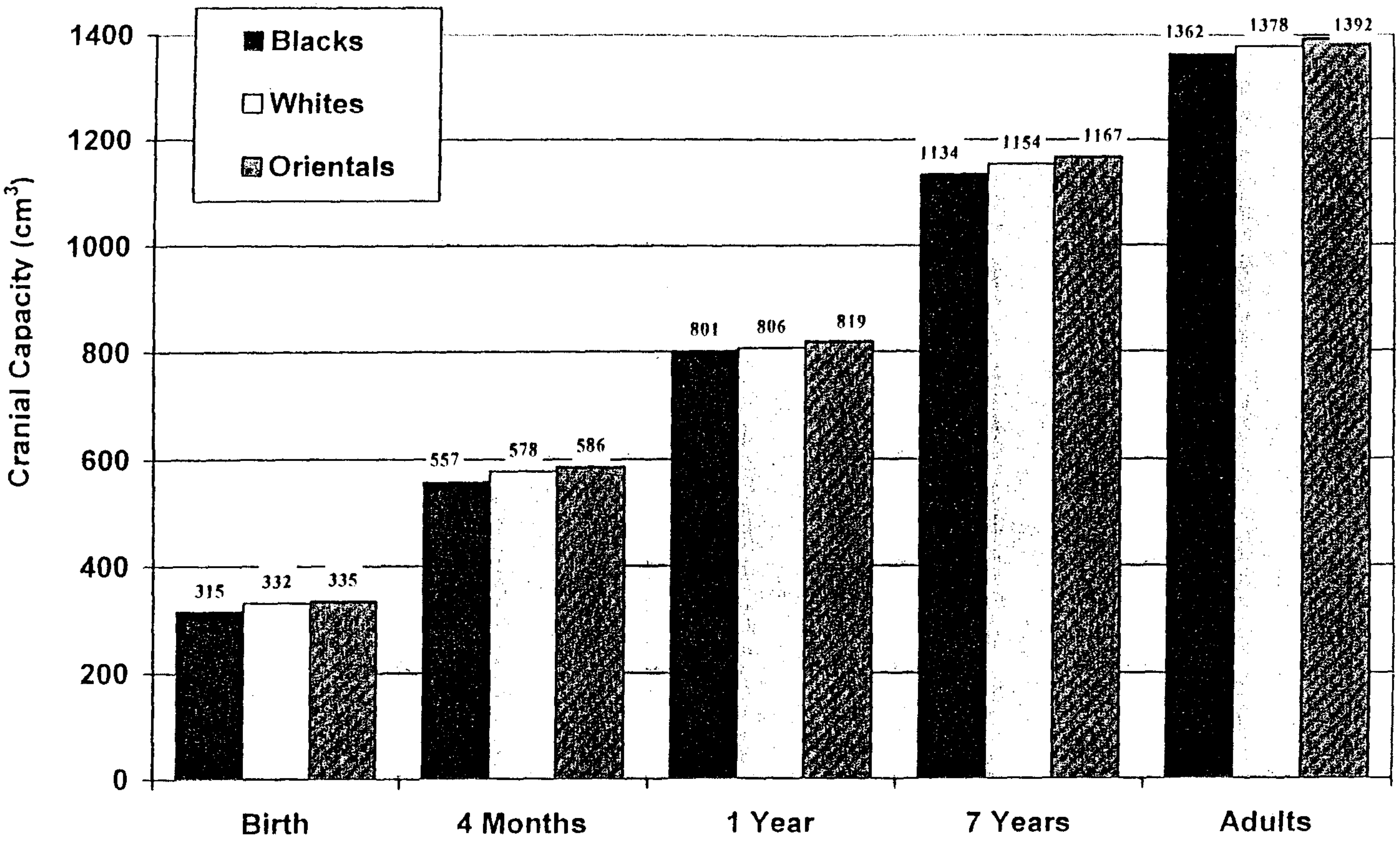

Figure 9.2: Average cranial capacity (cm^3) from birth to adulthood for Blacks, Whites, and Orientals in the U.S. Birth through age 7 from U.S. Collaborative Perinatal Project; adults from U.S. Army. (After Rushton 1997: 15).

demonstrated that even after adjusting for the effects of body size, sex and military rank in a stratified random sample of over 6,000 U.S. Army personnel, East Asians, Whites and Blacks averaged cranial capacities of 1,416, 1,380 and 1,359 cm^3, respectively. The third study (Rushton 1993) re-analyzed a set of anthropometric data originally published by Melville Herskovits who offered it as evidence against race differences in cranial capacity. The new analyses revealed that in fact Caucasoids averaged a cranial capacity of 1,421 and Negroids, 1,295 cm^3. The fourth study (Rushton 1994) analyzed data obtained on tens of thousands of people from around the world collated by the International Labour Office in Geneva and found that after adjusting for the effects of body size and sex, samples from the Pacific Rim, Europe, and Africa averaged cranial capacities, of 1,308, 1,297, and 1,241 cm^3, respectively. Finally, Rushton & Osborne (1995) analyzed the Georgia Twin Study of adolescents and found that after correcting for body size and sex, Whites averaged a cranial capacity of 1,269 and Blacks 1,251 cm^3.

Many are surprised to learn that the races differ in brain size (Kamin & Omari 1998; Lieberman 2001; Graves 2002), and they question how reliable the evidence is. In fact, dozens of studies from the 1840s to the 1990s, using different methods on different samples, reveal the same strong pattern. Three other methods of measuring brain size also reveal the same pattern of race differences: (1) endocranial volume from empty skulls; (2) wet brain weight at autopsy; and (3) high tech MRI. For example, using state-of-the-art MRI technology, Harvey *et al.* (1994) found that 41 Africans and West Indians averaged a smaller brain volume than did 67 Caucasians.

Using endocranial volume, the American anthropologist Samuel George Morton (1849) filled over 1,000 skulls with packing material and found that Blacks averaged about 5 cubic inches less cranial capacity than Whites. His results were confirmed by Todd (1923), Gordon (1934) and Simmons (1942). In 1984, Beals *et al.* carried out the largest study of race differences in endocranial volume to date, using 20,000 skulls from around the world. They reported that East Asians, Europeans, and Africans averaged cranial volumes of 1,415, 1,362, and 1,268 cm^3, respectively. The skulls from East Asia averaged 3 cubic inches larger than those from Europe, which in turn was 5 cubic inches larger than the African average.

Using the method of weighing brains at autopsy, the famous French neurologist Paul Broca (1873) found that Whites averaged heavier brains than did Blacks and showed more complex convolutions and larger frontal lobes. Broca corroborated the Black-White differences using the endocranial volume method as well as finding that East Asians averaged larger cranial capacities than did Europeans. These results too have stood the test of time. Subsequent autopsy studies have found an average Black-White brain weight difference of about 100 grams (Bean 1906; Mall 1909; Vint 1934; Pearl 1934). Some studies have found that the more White admixture (judged independently from skin color), the greater the average brain weight in Blacks (Bean 1906; Pearl 1934). A 1980 autopsy study of 1,261 American adults by Ho *et al.* (1980) found that 811 White Americans in this sample averaged 1,323 grams and 450 Black Americans averaged 1,223 grams — a difference of 100 grams. Since the Blacks and Whites in the study were similar in body size, it was not responsible for the differences in brain weight.

Rushton (1995) summarized the world database using the three methods on which there are a sufficient number of studies (autopsies, endocranial volume, head measurements) as well as head measurements corrected for body size (see pp. 126–132, Table 6.6). The results in cm^3 or equivalents were: East Asians = 1,351, 1,415, 1,335, 1,356 (mean = 1,364); Whites = 1,356, 1,362, 1,341, 1,329 (mean = 1,347); and Blacks = 1,223, 1,268, 1,284, and 1,294 (mean = 1,267). The overall mean for Asians is 17 cm^3 more than that for Europeans and 97 cm^3 more than that for Africans. Within-race differences, due to method of estimation, averaged 31 cm^3. Since one cubic inch of brain matter contains millions of brain cells and hundreds of millions of synapses or neural connections, these brain size differences help to explain why the races differ in average IQ.

As a committed Jensenist, I pursued these and many other hypotheses with vigor. To account for the trade-off between racial differences in brain size and egg-production and all the other traits shown in Table 9.4, Rushton (1988, 1995, 2000) proposed a gene-based "life-history theory" based on evolutionary biology's $r - K$ scale of reproductive strategies. This scale is generally used to compare the life histories of widely disparate species, but Rushton used it to describe the smaller but real differences between the human races. East Asians are more K and so tend to devote resources to producing small numbers of children and invest heavily in them and provide them with a high level of parental care; Africans are more r and devote resources to producing greater numbers of children, invest less heavily in them and give them less parental care; Whites are intermediate, though closer to East Asians. The r/K scale predicted a wide spectrum of characteristics including fertility, infant mortality, rate of physical maturation, intelligence, brain size, dizygotic twinning, crime, sexual potency, sexual precocity, number of sexual partners, and hormone levels. Highly K-selected women produce fewer eggs (and have bigger brains) than r-selected women. Highly K-selected men invest time and energy in their children rather than the pursuit of sexual thrills. They are "dads" rather than "cads".

Rushton (1988, 1995, 2000) also mapped the $r - K$ theory of racial differences onto the "Out of Africa" theory of human origins. Only when *Homo sapiens* left Africa, about 100,000 years ago, did they begin to develop the racial traits we see today, by adapting to the new regions and climates. The first major split was between the Africans and the non-Africans. Then about 40,000 years ago there was another major split, between the ancestors of today's Whites and East Asians. The African/non-African split occurred more than twice as early as the East Asian-White split. This explains why Whites average between East Asians and Africans on so many life history traits.

The climate differences influenced mental abilities. In Africa, food and warmth were available all year round. As *Homo sapiens* moved out of Africa they faced an entirely new problem — cold winters. Gathering and storing food, providing shelter, making clothes, and raising children during these long winters were more mentally demanding tasks than those that humans had faced previously. These tasks called for larger brains and slower rates of growth. They resulted in lower levels of sex hormones leading to fewer twins, less sexual behavior and aggression, and more family stability. Both parents had to provide more care to help their young survive in the harsher climates. Thus came about the pattern of traits in Table 9.4.

3. Genes? Environment? Or Both?

All the issues concerning Black-White differences in IQ that Jensen (1969a) raised in his famous *Harvard Educational Review* article are still with us today. Indeed, much of the opposition to IQ testing and heritability would probably disappear if it were not for the stubborn and unwelcome fact that, despite extensive well-funded programs of intervention, the Black-White difference refuses to go quietly into the night.

Jensen's long intellectual march on this topic led triumphantly to his latest book, *The g Factor* (1998). Jensen's tome does not draw back from Jensenist conclusions — that the average difference in IQ found between Blacks and Whites has a substantial hereditary component, that this difference is related mainly to the *g*-factor, and that it has important societal consequences. Jensen (1998: 418) proposed the "default hypothesis" for Black-White IQ differences, viz., that they are due to the same weight of genetic and environmental factors as are the causes of individual differences within each race. There is no need for any ad hoc hypothesis, or to postulate some Factor X, that is unique to either Blacks or Whites.

Chapter 12 of *The g Factor* presents Jensen's technical arguments for why he believes that Black-White IQ differences are about 50% genetic in origin. These include that: (1) the Black-White IQ differences are most pronounced on the more *g*-loaded components of IQ tests; (2) the Black-White IQ differences are most pronounced on the more heritable components of IQ tests; (3) IQ differences are associated with brain size within each race and there are significant Black-White (and East Asian) differences in average brain size; (4) Black-White (and East Asian) differences show up in myopia which has been linked to brain size; (5) the Black-White (and East Asian) IQ differences remain following transracial adoption; (6) the Black-White IQ differences are reflected in studies of racial admixture; (7) the Black-White IQ differences are predicted by "regression to the mean"; (8) Black-White-East Asian differences in neonate behavior, rate of maturation, and a suite of life-history traits parallel the IQ differences; (9) the Black-White-East Asian IQ differences cannot be explained by any culture-only theory; and (10) the Black-White-East Asian IQ differences dovetail with what is known about human evolution. What follows is a summary of some of the evidence from Jensen's (1998) The *g* Factor and Rushton's (1995, 2000) *Race, Evolution, and Behavior.*

4. Evidence for the Default Hypothesis for Black-White IQ Differences

4.1. Black-White IQ Differences are Most Pronounced on the More g-Loaded Components of IQ Tests

As reviewed early in this chapter, Black-White differences are Jensen Effects, being most pronounced on the more *g*-loaded subtests. Spearman's hypothesis thus constitutes a special case of the Jensen Effect. It applies even to the *g* factor extracted from reaction time measures taken from 9- to 12-year-old Black and White children. Jensen (1998) has shown that a test's *g* loading is the best predictor not just of that test's correlation with scholastic and work-place performance, but of heritability coefficients determined from twin studies, inbreeding depression scores calculated in children born from cousin-marriages, and many other variables including brain evoked potentials, brain pH levels, brain glucose metabolism, as well as nerve conduction velocity, reaction time and other physiological factors. These correlations establish the heritable and biological, as opposed to the mere statistical, reality of *g*. The general factor *g* is a product of human evolution. Indeed, massive evidence indicates that *g* is related to the size and functioning of the brain (Duncan *et al.* 2000; Rushton & Ankney 1996; see also Chapters 6 and 10 in this volume). As reviewed above, race differences in brain size occur at birth and continue through life.

4.2. Black-White IQ Differences are Most Pronounced on the More Heritable Components of IQ Tests

Individual differences are heritable *within* races, indeed within all species and sub-species so far studied. Dozens of twin, adoption, and family studies have confirmed the high heritabilities for intellectual and social variables within human races (as reviewed by Bouchard 1996; Bouchard & Loehlin 2001 and Chapter 7 of *The g Factor*). By simple generalization, therefore, we would expect that race differences are heritable too. If, however, environmental deprivation is stronger for Blacks than for Whites, the heritabilities for Blacks should be reduced. If so, greater environmental *damage* and not genes would be the cause of the race difference.

Loehlin *et al.* (1975: 114–116) reviewed the literature to date and found that while there was some evidence suggesting a lower heritability of intelligence for Blacks than for Whites (e.g. by Scarr-Salapatek 1971), a larger body of evidence suggested equal heritabilities in the two groups. Subsequently, Osborne's (1980b) Georgia Twin Study compared 123 Black and 304 White pairs of 12- to 18-year-old twins drawn from schools in Georgia, Kentucky, and Indiana, given the Basic Test Battery, along with smaller sub-sets of twins given the Primary Mental Abilities test and the Cattell Culture Fair Intelligence test. He found heritabilities of about 50% for both Blacks and Whites, all significantly different from zero but not from each other. (The heritabilities of the Basic, Primary, and Cattell tests respectively were, for Whites: 0.61, 0.37, and 0.71; and for Blacks: 0.75, 0.42, and 0.19; Osborne 1980b, pp. 68–69, 89, 98). Moreover, the heritabilities increased with age in Blacks, just as they did in Whites, indicating no evidence for the cumulative environmental deficits predicted by culture-only theory (Osborne 1980b, ch. XI).

Jensen (1998: 465) re-analyzed the Georgia Twin Study using structural equation modeling. This decomposes a phenotypic mean difference into its genetic and environmental components. Essentially, this methodology is a multiple regression technique that tests the "goodness-of-fit" of different alternative models that explain whether a difference between groups is due to the same genetic and environmental factors that cause individual differences within the groups, *or* whether some additional, minority-specific, cultural factor (an unknown Factor X) causes differences *between* groups but not differences *within* groups. Jensen (1998) tested three alternative models — *only* genetic factors, *only* environmental factors, or *neither* genes *nor* environment — against the default model (genes and environment). He found that the observed Black-White differences were best explained by both genetic and environmental factors, while either genetic or environmental explanations alone were inadequate.

Others too have used structural equation models to examine the genetic and cultural contributions to race differences. In a series of studies, Rowe (1994; Rowe *et al.* 1994, 1995) analyzed diverse but representative data sets. In one study of six data sources, Rowe *et al.* (1994) compared cross-sectional correlation matrices (about 10 × 10) for a total of 8,528 Whites, 3,392 Blacks, 1,766 Hispanics and 906 Asians. These matrices contained both independent variables (e.g. home environment, peer characteristics) and developmental outcomes (e.g. achievement, delinquency). When the matrices were compared by a LISREL goodness-of-fit test, each ethnic group's covariance matrix was equal to the matrix of the other groups. Not only were these matrices nearly identical but also they were no less alike than covariance matrices computed from random halves within one ethnic or racial group. There were no distortions in the matrices that required any minority-specific developmental Factor X to explain the correlations between the background variables and the outcome measures.

In another study, Rowe *et al.* (1995) extended this cross-sectional line of research by examining longitudinal data on academic achievement. Once again, the existence of any minority-specific cultural processes affecting achievement should produce different covariance structures among ethnic and racial groups. Correlation matrices were computed on academic achievement and family environment measures in 565 full-sibling pairs from the National Longitudinal Survey of Youth, each tested at ages 6.6 and 9.0 years (White $N = 296$ pairs; Black $N = 149$ pairs; Hispanic $N = 120$ pairs). Each population group was treated separately, yielding three 8 × 8 correlation matrices. When compared employing a LISREL method, the matrices were equal across the three groups. As a single structural equation model fitted all groups, the hypothesis of special minority-specific developmental processes affecting academic achievement was not supported.

Subsequently, Rowe & Cleveland (1996) extended the structural equation modeling studies to estimate explicitly the within-race heritabilities from Black and White full- and half-siblings, again with data from the National Longitudinal Survey of Youth (106 pairs of Black half-sibs, 53 pairs of White half-sibs; 161 pairs of Black full-sibs, 314 pairs of White full-sibs). Three Peabody Individual Achievement Tests were used (Mathematics, Reading Comprehension and Reading Recognition). The data fit the default hereditarian model that the sources of individual differences and of differences between racial means were the same — about 50% genetic and 50% environmental — extremely well.

Large-scale studies of military samples have also reported a nearly identical statistical structure on intellectual variables across races. Ree & Carretta (1995) examined a nationally representative sample of young Black, White and Hispanic men and women who took the Armed Services Vocational Aptitude Battery (ASVAB; $N = 9{,}173$). The ASVAB, which is used to select applicants for all military enlistments and assign them to first jobs, consists of 10 separately scored sub-tests (General Science, Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Numerical Operations, Coding Speed, Auto and Shop Information, Mathematics Knowledge, Mechanical Comprehension, Electronics Information). Ree and Carretta found the hierarchical factor structure of ASVAB sub-test scores was virtually identical across the three groups. Similarly, Carretta & Ree (1995) examined the more specialized and diverse Air Force Officer Qualifying Test (AFOQT), a multiple-aptitude battery that had been given to 269,968 applicants (212,238 Whites, 32,798 Blacks, 12,647 Hispanics, 9,460 Asian Americans and 2,551 Native Americans). Hierarchical *g* accounted for the greatest amount of variance in all groups and its loadings differed little by ethnicity. Thus, the factor structure of cognitive ability is nearly identical for Blacks and for Whites. These findings are consistent with the default hereditarian hypothesis.

Heritability data are especially informative when genetic theory and culture-only theories of race differences make diametrically opposite predictions. For example, genetic theory predicts that race differences will be greater on those sub-tests that are more heritable within races, while culture-only theory predicts that race differences will be greater on those sub-tests that are culturally malleable (i.e. those with lower heritability) and on which races can grow apart as a result of dissimilar experiences. Analyses of independent data sets support the genetic hypothesis.

Jensen (1973, ch. 4) was one of the first to apply differential heritabilities to the study of race differences. He calculated the *environmentality* of tests in both Black and White children defined as the degree to which sibling correlations departed from the pure genetic expectation of 0.50. Environmentality was inversely related to the magnitude of the Black-White difference ($r = -0.70$), leading to the conclusion that the more environmentally influenced a test, the less pronounced its Black-White difference. Jensen (1973) also cited an unpublished study by Nichols (1972) that estimated the heritability of 13 tests from 543 pairs of 7-year-old siblings, including an equal number of Blacks and Whites. Jensen found a 0.67 correlation between the heritability of a test and the magnitude of the Black-White difference on that test.

Prompted by Jensen's approach, Rushton (1989) estimated genetic weights by using the amount of inbreeding depression found on the 11 tests of the Wechsler Intelligence Scale for Children (WISC). Inbreeding depression occurs in the offspring of closely related parents when harmful recessive genes combine. The fact that inbreeding depression lowers the IQ in offspring in itself provides evidence for the heritability of IQ. Rushton found a positive correlation between inbreeding depression scores calculated from 1,854 cousin marriages in Japan and the magnitude of the Black-White difference in the U.S. on the same 11 Wechsler tests ($r = 0.48$; Figure 9.3). This contradicts culture-only theory, which predicts that differences between Blacks and Whites should be greater on those sub-tests most affected by the environment (i.e. those showing lowest amount of inbreeding depression). There really is no non-genetic explanation for the relation between inbreeding depression scores from Japan and Black-White differences in the U.S. Figure 9.3 also shows the regression of Black-White differences on the *g* factor (reviewed earlier). As either the *g* loadings or the inbreeding depression scores increase, the differences between Blacks and Whites also increase.

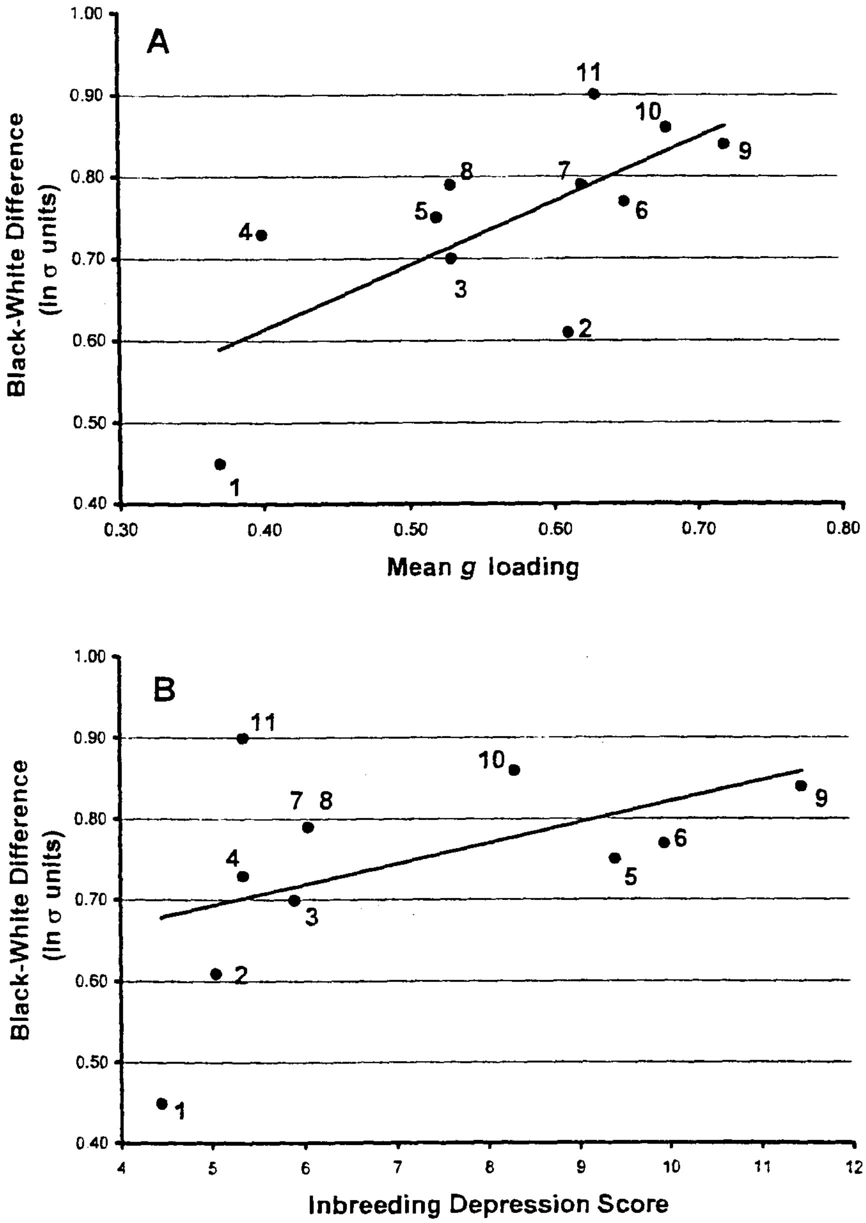

Figure 9.3: Regression of Black-White differences on *g* loadings (panel A) and on inbreeding depression scores (panel B). *Note:* The numbers indicate sub-tests from the Wechsler Intelligence Scale for Children—Revised: 1 Coding; 2 Arithmetic; 3 Picture completion; 4 Mazes; 5 Picture arrangement; 6 Similarities; 7 Comprehension; 8 Object assembly; 9 Vocabulary; 10 Information; 11 Block design. (After Rushton 1995: 188).

4.3. *Black-White (and East Asian) IQ Differences Occur in Average Brain Size Which is Linked to IQ scores Within Races*

The g Factor discusses Jensen's own studies on the relation between brain size and intelligence and to racial differences in brain size, as well as those reviewed earlier in this chapter. Jensen & Johnson (1994) showed that for Blacks, as for Whites, the head size × IQ correlation exists within-families as well as between-families, indicating the intrinsic or functional relationship mentioned earlier. Equally important is the fact that within each sex, Blacks and Whites fit the same regression line of head size on IQ. When Blacks and Whites are perfectly matched for true-score IQ (i.e. IQ corrected for measurement error), at either the Black mean or the White mean, the overall average White-Black difference in head circumference is virtually nil. (Matching Blacks and Whites for IQ eliminates the average difference in head size, but matching the groups on head size does not equalize their IQs. This is what one would expect if brain size is only one of a number of brain factors involved in IQ.)

In another analysis of the Georgia Twin Study, Jensen (1994) showed that the Black-White difference in head/brain size is also related to the magnitude of the Black-White difference in *g*. The correlation coefficient of each test with the head measurements was correlated with the magnitude of the Black-White difference on that test, thus forming two vectors. The column vector of test × head-size correlations correlated 0.51 ($P < 0.05$) with the vector of standardized White-Black differences on each of the tests.

The final piece of evidence that the race difference in brain size mediates the race difference in IQ comes from an "ecological correlation" (widely used in epidemiological research) of 0.998 between mean brain size measures and mean IQ scores across the three races (Jensen 1998: 443). Figure 9.4, which plots the regression of median IQ on mean cranial capacity is almost perfectly linear, with a Pearson $r = 0.998$. Mean cranial capacity for each of the three races accurately predicts their mean IQs.

4.4. *Black-White (and East Asian) IQ Differences Show up in Myopia Which has been Linked to Brain Size*

Myopia (near-sightedness) is positively correlated with IQ. The relationship appears to be pleiotropic, that is, a gene affecting one of the traits also has some effect on the other (Cohn *et al.* 1988). Further, there are significant racial and ethnic differences in the frequency of myopia, with the highest rates found in East Asians, the lowest rates among Africans and Europeans intermediate (Post 1982). Among Europeans, Jews have the highest rate of myopia, about twice that of gentiles and about on a par with that of Asians. Miller (1994) suggested that myopia is caused by extra myelinization in the eye and is similarly linked to brain size.

4.5. *Black-White (and East Asian) IQ Differences Remain Following Transracial Adoptions*

The g Factor also cites the evidence of transracial adoption studies. Three studies have been carried out on Korean and Vietnamese children adopted into White American and

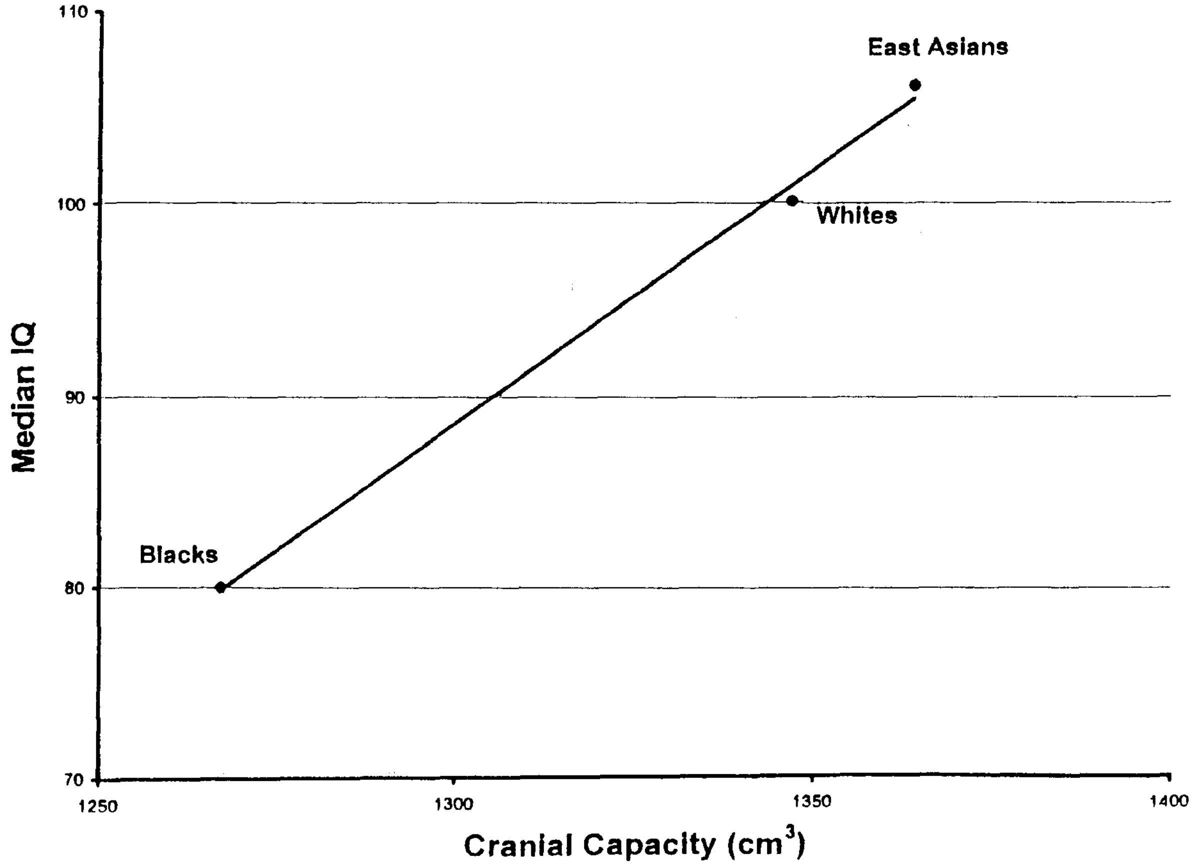

Figure 9.4: Median IQ of three populations (Mongoloid, Caucasoid, and Negroid) plotted as a function of the mean cranial capacity in each population. (Regression: $IQ = 0.262 \times \text{cranial capacity} - 252.6$; $r = 0.998$.) (After Jensen 1998: 443).

White Belgian homes (e.g. Frydman & Lynn 1989). Though many had been hospitalized for malnutrition prior to adoption, they went on to develop IQs ten or more points higher than their adoptive national norms. By contrast, Black and Mixed-Race (Black/White) children adopted into White middle-class families typically perform at a lower level than similarly adopted White children. The largest and best known such study, the Minnesota Transracial Adoption Study, was designed specifically by Sandra Scarr and Richard Weinberg to separate genetic factors from rearing conditions as causal influences on the poor cognitive performance of Black children (Scarr & Weinberg 1976; Weinberg *et al.* 1992). It is also the only transracial adoption study that includes a longitudinal follow-up, with testing at ages 7 and 17 years.

Scarr and her colleagues compared the IQ and academic achievement scores of Black, White, and mixed-race Black/White children adopted into privileged White families in Minnesota by adopting parents whose mean IQ was more than 1 SD above the population mean of 100 (see Table 9.5). The biological children of these parents were also tested. The first testing of 265 children was carried out in 1975 when they were 7 years old and the second in 1986 when the 196 remaining in the study were 17 years old. The 7-year-old White biological (non-adopted) children had an average IQ of 117 (Table 9.5, 2nd column), similar to that found for other children of upper-middle-class parents. The adopted children with two White biological parents had a mean IQ of 112. The adopted children with one Black and one White biological parent averaged 109. The adopted children with two Black biological parents had an average IQ of 97. (A mixed group of 21 Asian, North American Indian, and Latin American Indian adopted children averaged an IQ of 100 but were not included in the main statistical analyses.)

Scarr & Weinberg (1976) interpreted the results of their testing at age 7 as strong support for the culture-only interpretation of racial differences in intelligence. They drew special attention to the fact that the mean IQ of 107 for all "socially classified" Black children (i.e. those with either one or two Black parents) was significantly above the U.S. White mean. The poorer performance of children with two Black biological parents was attributed to their more difficult and later placement. Scarr and Weinberg also pointed out that this latter group had both natural and adoptive parents with somewhat lower educational levels and abilities (two points lower in adoptive parents IQ). They found no evidence for "the expectancy effects" hypothesis that adoptive parent' beliefs about the child's racial background influence the child's intellectual development. The mean score for 12 children wrongly believed by their adoptive parents to have two Black biological parents was virtually the same as for 56 children correctly classified by their adoptive parents as having one Black and one White biological parent.

Table 9.5 also presents the results for the 196 children retested at 17 years old (Weinberg *et al.* 1992). There were four independent assessments of the children's cognitive performance at this later age: (1) an individually administered IQ test; (2) an overall grade point average; (3) a class rank based on school performance; and (4) four special aptitude tests in school subjects administered by the educational authority, which Rushton (1995) averaged. The results are strikingly concordant with the earlier testing. The non-adopted White children had a mean IQ of 109, a grade point average of 3.0, a class rank at the 64th percentile, and an aptitude score at the 69th percentile. The

Table 9.5: Comparison of white biological and white, mixed-race, and black adopted children at age 7 and 17 raised in middle-class white families.

Children's background	**Age 7**	**Age 17**	**Age 17 Grade Point Average**	**Age 17 Class Rank (percentile)**	**Age 17 School Aptitude based on National Norms (weighted mean of 4 percentiles)**
Biological Parents	120	115			
Non-adopted, with 2 White biological parents (*N* at 7 = 143; *N* at 17 = 104)	116	109	3.0	64	69
Adopted, With 2 White biological parents (*N* at 7 = 25; *N* at 17 = 16)	112	106	2.8	54	59
Adopted, with 1 White, 1 Black biological parent (*N* at 7 = 68; *N* at 17 = 55)	109	99	2.2	40	53
Adopted, with 2 Black biological parents (*N* at 7 = 29; *N* at 17 = 21)	97	89	2.1	36	42

Adapted from *Race, Evolution, and Behavior* (Rushton 1995). Based on data from Weinberg *et al.* (1992).

adopted children with two White biological parents had a mean IQ of 106, a grade point average of 2.8, a class rank at the 54th percentile, and an aptitude score at the 59th percentile. The adopted children with one Black and one White biological parent had a mean IQ of 99, a grade point average of 2.2, a class rank at the 40th percentile, and an aptitude score at the 53rd percentile. The adopted children with two Black biological parents had a mean IQ of 89, a grade point average of 2.1, a class rank at the 36th percentile, and an aptitude score at the 42nd percentile. (The 12 remaining mixed group of Amerindian and Asian children had an IQ of 96 with no data provided on school achievement.)

Because different tests based on different standardization groups were used in the first testing than were used in the follow-up, the overall average difference of about eight IQ points (evident for all groups, including the non-adopted group) between the two test periods is of no theoretical importance for the hypothesis of interest. The relevant comparisons are those between the adopted groups *within* each age level. The mean of 89 for adopted children with two Black parents was slightly above the national Black mean of 85 but not above the Black mean for Minnesota.

4.6. Black-White IQ Differences are Reflected in Studies of Racial Admixture

In the Minnesota Transracial Adoption Study, the Mixed-Race (Black/White) adoptees had a mean IQ between those of the "non-mixed" White and "non-mixed" Black adoptees, as predicted from a genetic hypothesis (see Table 9.5). Although Jensen (1998: 478–483) himself is equivocal on this topic, in fact many other studies report similar results. For example, with respect to IQ scores, Shuey (1966) found that in 16 of 18 studies, Blacks with lighter skin color averaged higher IQ scores than did those with darker skin. Shockley (1973) estimated that for low IQ Black populations there is a one-point increase in average "genetic" IQ for each 1% of Caucasian ancestry, with diminishing returns as an IQ of 100 is reached. The genetic hypothesis is also consistent with the African American mean IQ of 85 being 15 points above the African average of 70 (reviewed earlier), given the approximately 20% White admixture in this group (Chakraborty *et al.* 1992; Parra *et al.* 1998). Corroborating data come from the mixed-race "Colored" population of South Africa showing they too have an average IQ of 85, intermediate to the "pure" Africans and "pure" Whites (Owen 1992). What brain weight data are available also fit with the genetic hypothesis. Both Bean (1906) and Pearl (1934) found that the greater the amount of White admixture (judged independently from skin color), the higher the average brain weight at autopsy in Blacks.

Most recently, Lynn (2002) and Rowe (2002) have analyzed data from large, publicly available, archival data sets, which show that groups of mixed-race individuals have mean scores intermediate to unmixed groups of Blacks and of Whites. Lynn examined the 1982 National Opinion Research Center's survey of a representative sample of the adult population, excluding non-English speakers. The 442 Blacks in the sample were asked whether they would describe themselves as "very dark", "dark brown", "medium brown", "light brown", or "very light". The correlation between these self-ratings and a 10-word vocabulary test score was $r = 0.17$ ($p < 0.01$). Rowe examined the 1994

National Longitudinal Study of Adolescent Health's survey of a representative sample of youths, with intentional over-sampling of Black children of highly educated parents. The mean age for the entire sample (9,830 Whites, 4,017 Blacks, and 119 mixed-race individuals) was 16. The Black adolescents averaged a lower birth weight, a lower verbal IQ, and a higher number of sexual partners than did the White adolescents. For each characteristic, the mixed-race mean fell between the means of the other two groups. Rowe found the social class explanation of the group differences "unconvincing", because of the three variables, only verbal IQ showed a moderate correlation with social class and statistically adjusting for it left the main findings unchanged. He also rejected the "discrimination based on skin tone" hypothesis since it was eliminated by deliberately selecting only those mixed-race adolescents who were judged by their interviewers to be Black based on their physical appearance.

4.7. Blacks and Whites Regress Toward their Predicted Racial Means

Regression toward the mean provides still another way to test if race differences are genetic. Regression toward the mean is seen, on average, when high IQ people mate and their children are less intelligent than their parents. This is because the parents pass on some, but not all, of their exceptional genes to their offspring. The converse happens for very low IQ parents; they have children with higher IQs. Although parents pass on a random half of their genes to their offspring, they cannot pass on the particular combinations of genes that cause their own exceptionality. It's like rolling a pair of dice and having them come up two sixes or two ones. The odds are that on the next roll, you'll get some value that is not quite as high (or as low). Physical and psychological traits involving dominant and recessive genes show some regression effect. Genetic theory predicts the magnitude of the regression effect to be lesser the closer the degree of kinship (e.g. identical twin > full-sibling > half-sibling). Culture-only theory makes no systematic or quantitative predictions based on genetic kinship *per se*.

For any trait, regression predicts that scores will move towards the average for that race. So in the United States, genetic theory predicts that the children of Black parents of IQ 115 will regress toward the Black IQ average of 85, while children of White parents of IQ 115 will regress toward the White IQ average of 100. There are similar predictions for the low end of the scale. Children of Black parents of IQ 70 should move up toward the Black IQ average of 85, while children of White parents of IQ 70 should move up toward the White IQ average of 100. Regression to the mean has been tested and the predictions proved to be true many times over. But more importantly, both the White and Black groups show the same degree of regression throughout the entire range of IQs between ±3 SDs from the group mean. The Law of Regression also explains why Black children born to high IQ, wealthy Black parents have test scores two to four points lower than do White children born to low IQ poor White parents (Jensen 1998: 358). The high IQ Black parents were unable to pass on their advantage to their children, even though they gave them a good upbringing and good schools. (The same phenomenon, of course, is also found for high IQ White parents.) Again, no culture-only theory *predicts* these results. To do so it would have to invoke the ad hoc hypothesis that cultural factors perfectly imitate the effect theoretically predicted by genetic theory and confirmed repeatedly in studies of physical traits and in animals.

Jensen (1973, Chapter 4) tested the regression predictions with data from siblings. These provide an even better test than parent-offspring comparisons because siblings share very similar environments. Black and White children matched for IQ had siblings who had regressed approximately halfway to their respective population means rather than to the mean of the combined population. For example, when Black and White children are matched with IQs of 120, the siblings of Black children average close to 100, while the siblings of White children average close to 110. A reverse effect is found with children matched at the lower end of the IQ scale. When Black and White children are matched for IQs of 70, the siblings of Black children average about 78, while the siblings of White children average about 85. The regression line showed no significant departure from linearity throughout the range of IQ from 50 to 150.

4.8. Black-White-East Asian Differences in Neonate Behavior, Rate of Maturation and a Suite of Life-History Traits Parallel the IQ Differences

On average, Black babies are born a week earlier than White babies, yet they are more mature as measured by bone development. In America, 51% of Black children have been born by week 39 of pregnancy compared with 33% of White children. In Europe, Black African babies, even those born to mothers in the professional classes, are born earlier than White babies. These Black babies are not born premature. They are born sooner, but biologically they are more mature. The length of pregnancy depends on the genes.

After birth, Black babies continue to mature faster than White babies, while East Asian babies mature more slowly. X-rays show that the bones grow faster in Black children than in White children and faster in Whites than in East Asians. Black babies also have greater muscular strength and can reach for objects better. Their neck muscles are often so developed that they can lift their heads up when they are only nine hours old. In a matter of days they can turn themselves over. Black children sit, crawl, walk, and put on their own clothes earlier than Whites or East Asians. East Asian children, on the other hand, mature more slowly than do White children. East Asian children often do not walk until 13 months while White children average walking at 12 months and Black children average walking at 11 months.

Blacks have faster dental development than do Whites, who mature faster than do East Asians. For example, Black children begin the first stage of permanent tooth growth at about 5.8 years while Whites and East Asians don't begin until 6.1 years. Blacks also reach sexual maturity sooner than do Whites, who in turn reach sexual maturity sooner than do East Asians. This is true for things like age of first menstruation, age of first sexual experience, and age of first pregnancy (Rushton 1995). It is unlikely that social factors could produce these differences. Across species a slower rate of development tends to go with greater brain size.

As reviewed earlier (Table 9.4), data from around the world on over 60 different variables including speed of maturation and longevity, personality and temperament, family stability and crime, sexual behavior and fertility, as well as intelligence and brain size, show East Asians and Africans consistently average at opposite ends of a continuum, with Europeans intermediate. Studies of personality show that Blacks are on average more extraverted, outgoing and uninhibited than Whites, who are in turn more extraverted, outgoing and uninhibited than Asians. These differences in personality may be reflected in international differences in rate of violent crime, as reported in the INTERPOL Yearbooks. Analyzes of these data throughout the 1980s and 1990s showed that African and Caribbean countries had double the rate of violent crime than that of European countries and three times that of countries in the Pacific Rim. For example, Rushton & Whitney (2002) averaged the rate of three of these violent crimes (murder, rape, and serious assault) per 100,000 population for the years 1984, 1990, and 1996 and found rates of 142, 74, and 43 for Blacks, Whites, and East Asians, respectively. Similarly, the matrifocal family pattern found disproportionately among African Americans, and often related to the crime statistics, is to be found in Britain, Canada, the Caribbean, and in South-of-Saharan Africa (Draper 1989).

Parallel race differences exist in average testosterone level. Studies show 3% to 19% more testosterone in Black college students and military veterans than in their White counterparts (Ellis & Nyborg 1992) and a lower amount of testosterone among the Japanese than among White Americans (Polednak 1989). Because testosterone is a sex hormone that travels everywhere throughout the body and affects many behavioral systems, it may be a "master switch" that sets the individual and the racial average position on an overall suite of characteristics. Testosterone level affects temperament, self-concept, aggression, altruism, crime and sexuality, in women as well as in men. Testosterone is also involved in secondary sexual characteristics such as muscularity and deepening of the voice.

Also associated with differences in sex hormones is the rate of double ovulation. For example, around the world, the rate of dizygotic twinning is less than 4 per 1,000 births among East Asians, 8 among Europeans, and 16 or greater among Africans (Table 9.4). Multiple birthing rates have been shown to be heritable. It is based on the race of the mother, regardless of the race of the father, as found in East Asian-European crosses in Hawaii and European-African crosses in Brazil (Bulmer 1970). Worldwide surveys also report higher average levels of sexual activity in Africans than in Europeans and especially in East Asians (Table 4). International fertility rates show the racial pattern. So do sexually transmitted disease rates within and between countries.

4.9. Black-White-East Asian Differences Cannot be Explained by Culture-Only Theory

When deciding whether genes are involved in the Black-White average IQ difference, or whether culture-only theory is correct, the following results should be considered. First, the mean difference in IQ scores has scarcely changed over the past 100 years (despite repeated claims that the gap is narrowing) and it can be observed as early as three years of age (Peoples *et al.* 1995). Controlling for overall socioeconomic level only reduces the mean difference by 4 IQ points. Culture-fair tests tend to give Blacks slightly *lower* scores, on the average, than do more conventional tests, as do non-verbal tests compared with verbal tests, and abstract reasoning tests compared with tests of acquired knowledge. Also, the Black-White differences show up on the *g*-factor extracted from culture-fair reaction time tests. The pattern of race differences shown in Table 9.4 is consistent across time and nation. Environmental explanations must account for *all* these differences — in IQ, brain size, myopia, speed of dental development, age of sexual maturity, testosterone level and number of multiple births. Genetic theory provides a single parsimonious explanation for all of them.

4.10. Black-White-East Asian Differences Map Onto Genetic Distance Measures and Dovetail with what is Known About Human Evolution

Finally, race differences can be examined from an evolutionary perspective to explain the worldwide clustering of traits. Jensen accepts the "Out-of-Africa" theory, that *Homo sapiens* arose in Africa about 200,000 years ago, expanded northwards beyond Africa about 140,000 years ago, with a European/East Asian split about 41,000 years ago (Stringer & McKie 1996). Evolutionary selection pressures were different in the hot savanna where Africans evolved than in the cold northern regions where Europeans evolved, or in the even colder Arctic regions where East Asians evolved. These ecological differences had not only morphologic but also behavioral consequences. Rushton (1995) proposed that the farther north the populations migrated out of Africa, the more they encountered the cognitively demanding problems of gathering and storing food, gaining shelter, making clothes and raising children during prolonged winters. As these populations evolved into present-day Europeans and East Asians, they underwent selective pressure for larger brains.

It is in this evolutionary context that Jensen (1998: 420–437) takes on the "race is a myth" brigade. As *Homo sapiens* migrated further away from Africa the random genetic mutations that occur at a constant rate in all living species accumulated, along with the adaptive changes. The resulting differences in allele frequencies are sufficient to warrant the designation of subspecies. Virtually every living species on earth has two or more subspecies. The human species is no exception, but then the subspecies are called races. Numerous and extensive genetic investigations yield essentially the same picture and identify the same major racial groupings as did the morphological markers of classical anthropology. The genetic evidence shows that, by far, the greatest divergence within the human species is between Africans (who have had the most time for random mutations to accumulate) and non-Africans (Cavalli-Sforza *et al.* 1994; Nei & Roychoudhury 1993). In a long footnote, Jensen (1998: 517–520) carried out a principal components analysis of data on genetic markers from Nei and Roychoudhury (1993) and found the familiar clustering of races: (1) Mongoloids; (2) Caucasoids; (3) South Asians and Pacific Islanders; (4) Negroids; (5) North and South Amerindians and Eskimos; and (6) Aboriginal Australians and Papuan New Guineans. Anyone wanting to argue, "race is only skin deep" has to confront the consistency of such results.

5. Conclusion

Most pieces of the scientific puzzle for why Blacks average lower IQ scores than do Whites are now falling into place. For example, the conclusion that intelligence is related to brain size and that there are racial differences in brain size, is becoming accepted. Ulric Neisser, Chair of the recent American Psychological Association's Task Force Report on *The Bell Curve* (Neisser *et al.* 1996), acknowledged that, with respect to "racial differences in the mean measured sizes of skulls and brains (with East Asians having the largest, followed by Whites and then Blacks) . . . there is indeed a small overall trend" (Neisser 1997: 80). Moreover, the average Black-White differences are now established using independent data sets and different test instruments around the world, in southern Africa and in the Netherlands, as in the United States. All are Jensen Effects. As Spearman (1927) predicted, those sub-tests that show the most pronounced Black-White differences are typically the ones with the highest g-loadings.

It is important to know that the Spearman-Jensen hypothesis is robust and that g is the same in southern Africa and the Netherlands as it is in the U.S. This tells us that the largest single source of Black-White differences around the world is essentially the same as the source of differences between individuals *within* each racial group — namely, g. This implies that a scientific understanding of Black-White, indeed of many individual, group and developmental differences, depends on understanding the nature of g. Race differences are not due to idiosyncratic cultural peculiarities in this or that test but to a general factor that all the ability tests measure in common.

Jensen's default hypothesis views mean population differences in g simply as aggregated individual differences and they are explainable by the same principles, thereby not violating Occam's razor by invoking unnecessary ad hoc hypotheses. Jensen's hypothesis is consistent with a preponderance of psychometric, behavior-genetic and evolutionary lines of evidence. And like true scientific hypotheses generally, it continually invites empirical refutation.

Jensen's methodological and theoretical analyses have distilled the deep essence of intelligence. Jensen has gone beyond proving the statistical reality and predictive validity of the general factor. He has shown Spearman's g to be a keystone of the behavioral sciences. If future psychometricians "see further", it will only be by standing on the shoulders of these two giants: Spearman and Jensen.

References

Avenant, T. J. (1988). *The establishment of an individual intelligence scale for adult South Africans. Report on an exploratory study conducted with WAIS-R on a sample of Blacks (Report No. P-91).* Pretoria, South Africa: Human Sciences Research Council.

Beals, K. L., Smith, C. L., & Dodd, S. M. (1984). Brain size, cranial morphology, climate, and time machines. *Current Anthropology, 25*, 301–330.

Bean, R. B. (1906). Some racial peculiarities of the Negro brain. *American Journal of Anatomy, 5*, 353–432.

Bouchard, T. J. Jr (1996). Behaviour genetic studies of intelligence, yesterday and today: The long journey from plausibility to proof. *Journal of Biosocial Science, 28*, 527–555.

Bouchard, T. J. Jr, & Loehlin, J. C. (2001). Genes, evolution, and personality. *Behavior Genetics*, *31*, 243–273.

Broca, P. (1873). Sur les crânes de la caverne de l'Homme Mort (Loere). *Revue d'Anthropologie*, *2*, 1–53.

Broman, S. H., Nichols, P. L., Shaughnessy, P., & Kennedy, W. (1987). *Retardation in young children*. Hillsdale, NJ: Erlbaum.

Bulmer, M. G. (1970). *The biology of twinning in man*. Oxford: Clarendon.

Carretta, T. R., & Ree, M. J. (1995). Near identity of cognitive structure in sex and ethnic groups. *Personality and Individual Differences*, *19*, 149–155.

Cavalli-Sforza, L. L., Menozzi, P., & Piazza, A. (1994). *The history and geography of human genes*. Princeton, NJ: Princeton University Press.

Centers for Disease Control (1988). Health status of Vietnam veterans. *Journal of the American Medical Association*, *259*, 2701–2719.

Chakraborty, R., Kamboh, M. I., Nwankwo, M., & Ferrell, R. E. (1992). Caucasian genes in American Blacks. *American Journal of Human Genetics*, *50*, 145–155.

Cohn, S. J., Cohn, C. M. G., & Jensen, A. R. (1988). Myopia and intelligence: A pleiotropic relationship? *Human Genetics*, *80*, 53–58.

Colom, R., Abad, F. J., Garcia, L. F., & Juan-Espinosa, M. (2002). Education, Wechsler's full-scale IQ, and *g*. *Intelligence*, *30*, 449–462.

Colom, R., Juan-Espinosa, M., & Garcia, L. F. (2001). The secular increase in test scores is a "Jensen effect". *Personality and Individual Differences*, *30*, 553–559.

Draper, P. (1989). African marriage systems: Perspectives from evolutionary ecology. *Ethology and Sociobiology*, *10*, 145–169.

Duncan, J., Seitz, R. J., Kolodny, J., Bor, D., Herzog, H., Ahmed, A., Newell, F. N., & Emslie, H. (2000). A neural basis for general intelligence. *Science*, *289*, 457–460.

Ellis, L., & Nyborg, H. (1992). Racial/ethnic variations in male testosterone levels: A probable contributor to group differences in health. *Steroids*, *57*, 72–75.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, *95*, 29–51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*, 171–191.

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist*, *54*, 5–20.

Frearson, W., & Eysenck, H. J. (1986). Intelligence, reaction time (RT) and a new "odd-man-out" RT paradigm. *Personality and Individual Differences*, *7*, 807–817.

Frydman, M., & Lynn, R. (1989). The intelligence of Korean children adopted in Belgium. *Personality and Individual Differences*, *10*, 1323–1326.

Gordon, H. L. (1934). Amentia in the East African. *Eugenics Review*, *25*, 223–235.

Gottfredson, L. S. (Ed.) (1997). Intelligence and social policy [Special Issue]. *Intelligence*, *24*, 1–320.

Graves, J. L. Jr (2002). The misuse of life-history theory: J. P. Rushton and the pseudoscience of racial hierarchy. In: J. L. Fish (Ed.), *Race and intelligence: Separating science from myth* (pp. 57–94). Mahwah, NJ: Erlbaum.

Grieve, K. W., & Viljoen, S. (2000). An exploratory study of the use of the Austin Maze in South Africa. *South African Journal of Psychology*, *30*, 14–18.

Harvey, I., Persaud, R., Ron, M. A., Baker, G., & Murray, R. M. (1994). Volumetric MRI measurements in bipolars compared with schizophrenics and healthy controls. *Psychological Medicine*, *24*, 689–699.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve*. New York: Free Press.

Ho, K. C., Roessmann, U., Straumfjord, J. V., & Monroe, G. (1980). Analysis of brain weight: I & II. *Archives of Pathology and Laboratory Medicine*, *104*, 635–645.

Jensen, A. R. (1968). Patterns of mental ability and socioeconomic status. *Proceedings of the National Academy of Sciences*, *60*, 1330–1337.

Jensen, A. R. (1969a). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, *39*, 1–123.

Jensen, A. R. (1969b). Reducing the heredity-environment uncertainty. *Harvard Educational Review*, *39*, 449–483.

Jensen, A. R. (1973). *Educability and group differences*. London: Methuen.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Jensen, A. R. (1984). Sociobiology and differential psychology: The arduous climb from plausibility to proof. In: L. J. R. Royce, & L. P. Mos (Eds), *Annals of theoretical psychology* (Vol. 2, pp. 49–58). New York: Plenum.

Jensen, A. R. (1985). The nature of the black-white difference on various psychometric tests: Spearman's hypothesis. *Behavioral and Brain Sciences*, *8*, 193–263.

Jensen, A. R. (1987). Further evidence for Spearman's hypothesis concerning the black-white differences on psychometric tests. *Behavioral and Brain Sciences*, *10*, 512–519.

Jensen, A. R. (1993). Spearman's hypothesis tested with chronometric information-processing tasks. *Intelligence*, *17*, 47–77.

Jensen, A. R. (1994). Psychometric *g* related to differences in head size. *Personality and Individual Differences*, *17*, 597–606.

Jensen, A. R. (1998). *The g factor*. Westport, CT: Praeger.

Jensen, A. R. (in press). Do age-group differences on mental tests imitate racial differences? *Intelligence*.

Jensen, A. R., & Johnson, F. W. (1994). Race and sex differences in head size and IQ. *Intelligence*, *18*, 309–333.

Jensen, A. R., & Whang, P. A. (1993). Reaction times and intelligence: A comparison of Chinese-American and Anglo-American children. *Journal of Biosocial Science*, *25*, 397–410.

Jensen, A. R., & Whang, P. A. (1994). Speed of accessing arithmetic facts in long-term memory: A comparison of Chinese-American and Anglo-American children. *Contemporary Educational Psychology*, *19*, 1–12.

Kamin, L., & Omari, S. (1998). Race, head size, and intelligence. *South African Journal of Psychology*, *28*, 119–128.

Lieberman, L. (2001). How 'Caucasoids' got such big crania and why they shrank: From Morton to Rushton. *Current Anthropology*, *42*, 69–95.

Loehlin, J. C., Lindzey, G., & Spuhler, J. N. (1975). *Race differences in intelligence*. San Francisco, CA: Freeman.

Lynn, R. (1977). The intelligence of the Japanese. *Bulletin of the British Psychological Society*, *30*, 69–72.

Lynn, R. (1978). Ethnic and racial differences in intelligence: International comparisons. In: R. T. Osborne, C. E. Noble, & N. Weyl (Eds), *Human variation: The biopsychology of age, race, and sex* (pp. 261–286). New York, Academic.

Lynn, R. (1982). IQ in Japan and the United States shows a growing disparity. *Nature*, *297*, 222–223.

Lynn, R. (1991). Race differences in intelligence: A global perspective. *Mankind Quarterly*, *31*, 255–296.

Lynn, R. (1997). Geographical variation in intelligence. In: H. Nyborg (Ed.), *The scientific study of human nature: Tribute to Hans J. Eysenck at eighty*. London: Elsevier.

Lynn, R. (2002). Skin color and intelligence in African Americans. *Population and Environment*, *23*, 365–375.

Lynn, R. (2003). The geography of *g*. In: H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen.* London: Elsevier.

Lynn, R., & Owen, K. (1994). Spearman's hypothesis and test score differences between Whites, Indians, and Blacks in South Africa. *Journal of General Psychology, 121*, 27–36.

Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations.* Westport, CT: Praeger.

Mall, F. P. (1909). On several anatomical characters of the human brain, said to vary according to race and sex, with special reference to the weight of the frontal lobe. *American Journal of Anatomy, 9*, 1–32.

Miller, E. M. (1994). Intelligence and brain myelination: A hypothesis. *Personality and Individual Differences, 17*, 803–832.

Morton, S. G. (1849). Observations on the size of the brain in various races and families of man. *Proceedings of the Academy of Natural Sciences Philadelphia, 4*, 221–224.

Must, O., Must, A., & Raudik, V. (in press (a)). The Flynn Effect for gains in literacy found in Estonia is not a Jensen Effect. *Personality and Individual Differences.*

Must, O., Must, A., & Raudik, V. (in press (b)). The secular rise in IQs: In Estonia the Flynn Effect is not a Jensen Effect. *Intelligence.*

Nagoshi, C. T., Johnson, R. C., DeFries, J. C., Wilson, J. R., & Vandenberg, S. G. (1984). Group differences and first principal-component loadings in the Hawaii Family Study of Cognition: A test of the generality of "Spearman's hypothesis". *Personality and Individual Differences, 5*, 751–753.

Nei, M., & Roychoudhury, A. K. (1993). Evolutionary relationships of human populations on a global scale. *Molecular Biology and Evolution, 10*, 927–943.

Neisser, U. (1997). Never a dull moment. *American Psychologist, 52*, 79–81.

Neisser, U., Boodoo, G., Bouchard, T. J. Jr, Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 15*, 77–101.

Nell, V. (2000). *Cross-cultural neuropsychological assessment: Theory and practice.* Mahwah, NJ: Erlbaum.

Nichols, P. L. (1972). *The effects of heredity and environment on intelligence test performance in 4- and 7-year-old white and Negro sibling pairs.* Unpublished doctoral dissertation, University of Minnesota.

Nyborg, H., & Jensen, A. R. (2000). Black-White differences on various psychometric tests: Spearman's hypothesis tested on American armed services veterans. *Personality and Individual Differences, 28*, 593–599.

Osborne, R. T. (1980a). The Spearman-Jensen hypothesis. *Behavioral and Brain Sciences, 3*, 351.

Osborne, R. T. (1980b). *Twins: black and white.* Athens, GA: Foundation for Human Understanding.

Owen, K. (1992). The suitability of Raven's Standard Progressive Matrices for various groups in South Africa. *Personality and Individual Differences, 13*, 149–159.

Parra, E. J., Marcini, A., Akey, J., Martinson, J., Batzer, M. A., Cooper, R., Forrester, T., Allison, D. B., Deka, R., Ferrell, R. E., & Shriver, M. D. (1998). Estimating African American admixture proportions by use of population specific alleles. *American Journal of Human Genetics, 63*, 1839–1851.

Pearl, R. (1934). The weight of the Negro brain. *Science, 80*, 431–434.

Peoples, C. E., Fagan, J. F. III, & Drotar, D. (1995). The influence of race on 3-year-old children's performance on the Stanford-Binet (4th ed.). *Intelligence, 21*, 69–82.

Polednak, A. P. (1989). *Racial and ethnic differences in disease.* Oxford: Oxford University.

Post, R. H. (1982). Population differences in visual acuity: A review, with speculative notes on selection relaxation. *Social Biology, 29,* 319–343.

Raven, J., Summers, B., Birchfield, M., Brosier, G., Burciaga, L., Byrkit, B., *et al.* (1990). *Manual for Raven's progressive matrices and vocabulary scales. Research supplement No. 3: American and international norms* (2nd ed.). Oxford: Oxford Psychologists Press.

Raven, J. C., Court, J. H., & Raven, J. (1998). *Manual for Raven's standard progressive matrices.* Oxford, U.K.: Oxford Psychologists Press.

Ree, M. J., & Carretta, T. R. (1995). Group differences in aptitude factor structure on the ASVAB. *Educational and Psychological Measurement, 55,* 268–277.

Roth, P. L., Bevier, C. A., Bobko, P., Switzer III, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54,* 297–330.

Rowe, D. C. (1994). No more than skin deep. *American Psychologist, 49,* 215–216.

Rowe, D. C. (2002). IQ, birth weight, and number of sexual partners in White, African American, and mixed race adolescents. *Population and Environment, 23,* 513–524.

Rowe, D. C., & Cleveland, H. H. (1996). Academic achievement in Blacks and Whites: Are the developmental processes similar? *Intelligence, 23,* 205–228.

Rowe, D. C., Vazsonyi, A. T., & Flannery, D. J. (1994). No more than skin deep: Ethnic and racial similarity in developmental process. *Psychological Review, 101,* 396–413.

Rowe, D. C., Vazsonyi, A. T., & Flannery, D. J. (1995). Ethnic and racial similarity in developmental process: A study of academic achievement. *Psychological Science, 6,* 33–38.

Rushton, J. P. (1984a). Sociobiology: Toward a theory of individual and group differences in personality and social behavior. In: J. R. Royce, & L. P. Mos (Eds), *Annals of theoretical psychology* (Vol. 2, pp. 1–48). New York: Plenum.

Rushton, J. P. (1984b). Group differences, genetic similarity theory, and the importance of personality traits: Reply to commentators. In: J. R. Royce, & L. P. Mos (Eds), *Annals of theoretical psychology* (Vol. 2, pp. 73–81). New York: Plenum Press.

Rushton, J. P. (1988). Race differences in behaviour: A review and evolutionary analysis. *Personality and Individual Differences, 9,* 1009–1024.

Rushton, J. P. (1989). Japanese inbreeding depression scores: Predictors of cognitive differences between blacks and whites. *Intelligence, 13,* 43–51.

Rushton, J. P. (1991). Mongoloid-Caucasoid differences in brain size from military samples. *Intelligence, 15,* 351–359.

Rushton, J. P. (1992). Cranial capacity related to sex, rank, and race in a stratified random sample of 6,325 U.S. military personnel. *Intelligence, 16,* 401–413.

Rushton, J. P. (1993). Corrections to a paper on race and sex differences in brain size and intelligence. *Personality and Individual Differences, 15,* 229–231.

Rushton, J. P. (1994). Sex and race differences in cranial capacity from International Labour Office data. *Intelligence, 19,* 281–294.

Rushton, J. P. (1995). *Race, evolution, and behavior: A life history perspective.* New Brunswick, NJ: Transaction.

Rushton, J. P. (1997). Cranial size and IQ in Asian Americans from birth to age seven. *Intelligence, 25,* 7–20.

Rushton, J. P. (1998). The "Jensen Effect" and the "Spearman-Jensen Hypothesis" of Black-White IQ differences. *Intelligence, 26,* 217–225.

Rushton, J. P. (1999). Secular gains in IQ not related to the *g* factor and inbreeding depression — unlike Black-White differences: A reply to Flynn. *Personality and Individual Differences, 26,* 381–389.

Rushton, J. P. (2000). *Race, evolution, and behavior: A life history perspective* (3rd ed.). Port Huron, MI: Charles Darwin Research Institute.

Rushton, J. P. (2001). Black-White differences on the *g* factor in South Africa: A "Jensen Effect" on the Wechsler Intelligence Scale for Children-Revised. *Personality and Individual Differences*, *31*, 1227–1232.

Rushton, J. P. (2002). Jensen Effects and African/Colored/Indian/White differences on Raven's Standard Progressive Matrices in South Africa. *Personality and Individual Differences*, *33*, 65–70.

Rushton, J. P., & Ankney, C. D. (1996). Brain size and cognitive ability: Correlations with age, sex, social class and race. *Psychonomic Bulletin and Review*, *3*, 21–36.

Rushton, J. P., & Jensen, A. R. (in press). African-White IQ differences from Zimbabwe on the Wechsler Intelligence Scale for Children-Revised are mainly on the *g* factor. *Personality and Individual Differences*.

Rushton, J. P., & Osborne, R. T. (1995). Genetic and environmental contributions to cranial capacity in black and white adolescents. *Intelligence*, *20*, 1–13.

Rushton, J. P., & Skuy, M. (2000). Performance on Raven's Matrices by African and White university students in South Africa. *Intelligence*, *28*, 251–265.

Rushton, J. P., Skuy, M., & Fridjohn, P. (2002). Jensen effects among African, Indian, and White engineering students in South Africa on Raven's standard progressive matrices. *Intelligence*, *30*, 409–423.

Rushton, J. P., Skuy, M., & Fridjohn, P. (2003). Performance on Raven's advanced progressive Matrices by African, East Indian, and White engineering students in South Africa. *Intelligence*, *31*, 123–139.

Rushton, J. P., & Whitney, G. (2002). Geographic and populational variation in violent crime rates (from INTERPOL, 1993–1996). *Population and Environment*, *23*, 501–511.

Scarr-Salapatek, S. (1971). Race, social class and IQ. *Science*, *174*, 1285–1295.

Scarr, S., & Weinberg, R. A. (1976). IQ test performance of black children adopted by white families. *American Psychologist*, *31*, 726–739.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262–274.

Shockley, W. (1973). Variance of Caucasian admixture in Negro populations, pigmentation variability, and IQ. *Proceedings of the National Academy of Sciences, USA*, *70*, 2180a.

Shuey, A. M. (1966). *The testing of Negro intelligence*. New York: Social Science Press.

Simmons, K. (1942). Cranial capacities by both plastic and water techniques with cranial linear measurements of the Reserve Collection: white and Negro. *Human Biology*, *14*, 473–498.

Skuy, M., Schutte, E., Fridjhon, P., & O'Carroll, S. (2001). Suitability of published neuropsychological test norms for urban African secondary school students in South Africa. *Personality and Individual Differences*, *30*, 1413–1425.

Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.

Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New York: Macmillan.

Stringer, C., & McKie, R. (1996). *African exodus*. London: Cape.

te Nijenhuis, J. (1997). *Comparability of test scores for immigrants and majority group members in the Netherlands*. Unpublished doctoral dissertation, Vrije Universiteit, Amsterdam, The Netherlands.

te Nijenhuis, J., Evers, A., & Mur, J. P. (2000). Validity of the Differential Aptitude Test for the assessment of immigrant children. *Educational Psychology*, *20*, 99–115.

te Nijenhuis, J., & van der Flier, H. (1997). Comparability of GATB scores for immigrants and majority group members: Some Dutch findings. *Journal of Applied Psychology*, *82*, 675–687.

te Nijenhuis, J., & van der Flier, H. (2001). Group differences in mean intelligence for the Dutch and Third World immigrants. *Journal of Biosocial Science*, *33*, 469–475.

te Nijenhuis, J., Tolboom, E., Resing, W., & Bleichrodt, N. (in press). Does cultural background influence the intellectual performance of children from immigrant groups? Validity of the RAKIT intelligence test for immigrant children. *European Journal of Psychological Assessment*.

Todd, T. W. (1923). Cranial capacity and linear dimensions, in white and Negro. *American Journal of Physical Anthropology*, *6*, 97–194.

Vernon, P. A., & Jensen, A. R. (1984). Individual and group differences in intelligence and speed of information processing. *Personality and Individual Differences*, *10*, 573–576.

Vernon, P. E. (1982). *The abilities and achievements of Orientals in North America*. New York: Academic.

Vint, F. W. (1934). The brain of the Kenya native. *Journal of Anatomy*, *48*, 216–223.

Weinberg, R. A., Scarr, S., & Waldman, I. D. (1992). The Minnesota Transracial Adoption Study: A follow-up of IQ test performance at adolescence. *Intelligence*, *16*, 117–135.

Zaaiman, H., van der Flier, H., & Thijs, G. D. (2001). Dynamic testing in selection for an educational programme: Assessing South African performance on the Raven Progressive Matrices. *International Journal of Selection and Assessment*, *9*, 258–269.

Zindi, F. (1994). Differences in performance. *The Psychologist*, *7*, 549–552.