PERGAMON

Personality and Individual Differences 34 (2003) 177–183

PERSONALITY AND INDIVIDUAL DIFFERENCES

www.elsevier.com/locate/paid

African–White IQ differences from Zimbabwe on the Wechsler Intelligence Scale for Children-Revised are mainly on the *g* factor

J. Philippe Rushton[a,*], Arthur R. Jensen[b]

[a]*Department of Psychology, The University of Western Ontario, London, Ontario, N6A 5C2, Canada*
[b]*The University of California at Berkeley, USA*

Received 12 November 2001; received in revised form 3 May 2002; accepted 14 June 2002

Abstract

African–White differences on the sub-tests of the Wechsler Intelligence Scale for Children-Revised (WISC-R) in Zimbabwe are like the Black–White differences in the US in being positively associated with the sub-tests' *g* loadings (*g* being the general factor of intelligence). Published means and S.D.s for 204 12- to 14-year-old Zimbabweans on 10 WISC-R sub-tests were compared against those for 1868 White Americans. A principal factor analysis of the correlation matrix from the US standardization sample for Whites, along with the point-biserial correlation of African–White standardized differences, showed that fully 77% of the between-group race variance was attributable to a single source, namely *g*.

Keywords: IQ scores; *g* factor; Jensen effect; Black–White differences

Sub-saharan Africans, on average, score two standard deviations (S.D.s) lower on IQ tests than do people of European descent, while in the US, African Americans (Blacks) score just over one S.D. lower than do Whites (Jensen, 1998; Lynn & Vanhanen, 2002; Rushton & Skuy, 2000). In the US, it is known that Black–White differences are more pronounced on more highly *g*-loaded tests than they are on less *g*-loaded tests, *g* being the general factor of intelligence. This relation between the magnitude of a sub-tests' *g*-loading and its relative Black–White difference is known as "Spearman's hypothesis," after the British psychologist Charles Spearman (1927, p. 379) because he was the first to suggest it. It has also been called the "Spearman-Jensen hypothesis", because it was Jensen (1985, 1987, 1998) who brought Spearman's hypothesis to widespread attention and who did the empirical work confirming it. More recently and more generally,

* Corresponding author. Tel.: +1-519-661-3685; fax: +1-519-850-2302.
E-mail address: rushton@uwo.ca (J. P. Rushton).

0191-8869/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0191-8869(02)00192-7

g-factor relationships have become known as "Jensen Effects", because otherwise there is no name for them, only a long explanation of how the effect is calculated. Jensen Effects are not omnipresent and their absence can be as informative as their presence. For example, Rushton (1999) found that the Flynn Effect is not a Jensen Effect because the secular rise in IQ does not appear to be on *g*.

The Black–White difference on the *g*-factor is the best-established Jensen Effect. In *The g Factor* (1998, chap. 11), Jensen summarized the results from 17 independent data sets from the US of nearly 45,000 Blacks and 245,000 Whites derived from 171 psychometric tests and showed *g* loadings consistently predicted the magnitude of the Black–White difference ($r=0.63$; Spearman $\rho=0.71$, $P<0.05$). The most recent test of Spearman's hypothesis, based on 19 cognitive tests given to US Army veterans (3335 Whites and 502 Blacks) showed a correlation of 0.72; $P<0.01$ (Nyborg & Jensen, 2000). Jensen Effects were borne out even among 3 year olds who were administered eight sub-tests of the Stanford–Binet, where the rank correlation between *g* loadings and the Black–White differences was 0.71 ($P<0.05$). Even when the *g* loading is calculated from performance on elementary reaction-time tasks, which correlate with IQ (such as moving the hand to press a button to turn off a light, which all children can do in less than 1 s), the correlations between the *g* loadings of these tasks and the Black–White differences range from 0.70 to 0.81.

Since the studies on which Jensen (1998) based his analysis were all carried out in the US, this might be thought a phenomenon of limited interest, with its explanation sought in local conditions. However, five studies from South Africa have found that Black–White IQ differences are mainly on the *g* factor. Lynn and Owen (1994) were the first to explicitly find this effect in their analysis of data from over 3000 African, White, and Indian high-school students given 10 subtests of the South African Junior Aptitude Test. They 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).

Subsequently, Rushton and 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. They 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, $r=0.34$ ($P<0.01$, $N=46$, $\rho=0.41$, $P<0.01$).

Rushton (2001) analyzed 10 sub-tests of the Weschler Intelligence Scale for Children-Revised (WISC-R) from data published by Skuy, Schutte, Fridjhon, and O'Carroll (2001) on 154 Black South African high school students from Johannesburg. The mean score for Whites was set at the US standardization sample mean of 10 (which included African Americans). The mean African–White differences were then calculated and also expressed in S.D. units, using the African S.D.s. 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, 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 ($\rho = 0.74$, 0.74, respectively, Ps < 0.005).

Rushton (in press) re-analyzed published data from Owen (1992) who had given the Standard Progressive Matrices in South Africa without time limits to 1056 White, 1063 Indian, 778 Colored, and 1093 Black 14 year olds. 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 Ps < 0.01). The effects remained regardless of the ethnic group from which the item *g*-loadings were taken.

Finally, Rushton, Skuy, and Fridjohn (in press) 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. 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 $\rho = 0.53$).

The present study aims to further reduce the uncertainty over whether, in Africa, the *g* factor is the primary source of Black–White differences. These Zimbabwean data provide the first sub-Saharan test of Spearman's hypothesis made outside of South Africa. Since the data were collected by an African educational psychologist this also reduces uncertainty about bias in the data collection.

1. Method

The first half of the data for this study are taken from an African Zimbabwean educational psychologist, Fred Zindi (1994), who carefully matched 204 Black Zimbabwean pupils and 202 White British children from London inner-city schools for age (12–14 years old), sex, and educational level, both samples being characterized as "working class". All students were tested individually and effort was made to form a well functioning and representative sample. On each of 10 sub-tests from the WISC-R, the Zimbabwean children scored about one SD below the British group. Across the 10 sub-tests, the Zimbabwean children had a total IQ score of 70, while the British children's total IQ score was 95, a difference of nearly two S.D.s. Although the Zimbabwean scores might have been depressed by a language factor since English was not the Zimbabwean children's first language, it could not have been by much because Zindi (1994) found almost the same magnitude of difference on the Raven's Standard Progressive Matrices, a much less culturally loaded test.

The second data set is from Jensen and Reynold's (1982) analyses of 1868 6- to 16-year-old White children from the national standardization sample for the WISC-R. It also contained approximately equal numbers of both sexes. Jensen and Reynolds obtained the means, SDs, and correlation matrices for the Whites directly from the test publishers, as the manual does not provide this information separately by race (Wechsler, 1974). The White sub-test correlations allow examining the hypothesis that the African–White differences are mainly on the *g* factor.

Note that in the present study, the absolute magnitude of the African–White difference is not the focus of interest, so it is unnecessary to carefully match the two samples. Rather the focus is on the relative weighting the various sub-tests contribute to any differences that do exist. Will African–White differences be on the (independently obtained) *g* factor? Or will they be more pronounced on some other factor, for example on the verbal but not on the performance factor? Or will there be no pattern and the differences only randomly (and meaninglessly) linked to the various sub-tests?

2. Results

Table 1 presents much of the data for analysis. It includes, by column: (1) the name of the WISC-R subtest; (2) the Black Zimbabwean means for the 10 sub-tests reported by Zindi (1994); (3) the Black Zimbabwean S.D.s for the 10 sub-tests reported by Zindi (1994); (4) the White American means for the various sub-tests reported by Jensen and Reynolds (1982); (5) the White American S.D.s for the various sub-tests reported by Jensen and Reynolds (1982); (6) the African–White mean difference for the various scaled sub-tests; (7) the African–White differences standardized (in S.D. units) using the geometric mean of the Black and White S.D.s; (8) the *g* loadings extracted from the correlation of these 10 subtests from the White WISC-R standardization data from Jensen and Reynolds (1982); and (9) the reliabilities of the sub-tests taken from the WISC-R manual.

Table 1
Means and S.D.s of 204 Black (B) Zimbabwean students and 1868 White (W) Americans on sub-tests of the Wechsler Intelligence Scale for Children-Revised

(1) WISC-R sub-test	(2) Black Mean	(3) Black S.D.	(4) White Mean	(5) White S.D.	(6) W–B scale difference	(7) Standardized W–B difference	(8) *g*	(9) Reliability
Information	7.44	2.02	10.41	2.91	2.97	1.23	0.74	0.85
Similarities	7.23	2.25	10.29	3.01	3.06	1.18	0.74	0.81
Arithmetic	7.43	2.15	10.37	2.84	2.94	1.19	0.59	0.77
Vocabulary	6.48	2.27	10.42	2.94	3.94	1.53	0.78	0.86
Comprehension	6.27	2.13	10.44	2.81	4.17	1.71	0.67	0.77
Picture completion	7.24	2.97	10.41	2.87	3.17	1.09	0.55	0.77
Picture arrangement	8.01	2.79	10.37	2.91	2.36	0.83	0.54	0.73
Block design	6.46	2.25	10.39	2.92	3.93	1.53	0.68	0.85
Object assembly	6.93	2.41	10.73	3.01	3.80	1.41	0.53	0.70
Coding	6.24	3.03	10.22	3.30	3.98	1.26	0.37	0.72

Geometric mean of Black and White S.D. = $(S.D._W)(S.D._B)$.

A first test of Spearman's hypothesis was performed by correlating the vector of g loadings (column 8) with the standardized African–White differences (column 7). The result was in the expected direction ($r = 0.36$, ns). This correlation is lower than those found in American samples, even though the overall effect size for race is only about half as large in the US as it is in Africa. Unfortunately, the African sub-test correlations are not available to us to test any hypotheses about this apparent anomaly. However, additional analyses showed this correlation is raised to marginal significance ($r = 0.44$, $\rho = 0.52$; $P < 0.10$) when the African S.D.s (column 3) are used to calculate the African-White effect sizes and if these are then correlated with an *aggregated* set of WISC-R g loadings (as in Rushton, 1999).

As the essence of Spearman's hypothesis is that the Black–White difference is primarily a difference in g rather than in any other factors independent of g, a more direct method for testing it was made. Following Jensen (1987, p. 514, Table 1), we carried out a principal factor analysis on the correlation matrix (Table 2) for the US standardization sample for Whites (from Jensen & Reynolds, 1982), along with the effect sizes entered of the African–White differences as measured by point-biserial correlations from a formula given by Jensen (1980, p. 122). The positive correlation coefficients indicate that the Whites scored higher than the Africans.

Table 3 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 of g variance to all non-g variance is 3.5 to 1. Indeed, the race variable has a larger g loading than does any of the sub-test variables. It appears to reflect almost pure g.

3. Discussion

The main purpose of the study was to test whether the g factor in the US WISC-R standardization data would predict Black–White differences in a data set from Zimbabwe in southern

Table 2
Intercorrelations among WISC-R sub-tests for Whites plus the American White–Zimbabwean Black race[a] variable

Subtest	I	S	A	V	C	PC	PA	BD	OA	Cod	Race
Information	1	0.58	0.51	0.66	0.51	0.35	0.37	0.44	0.34	0.26	0.52
Similarities		1	0.43	0.63	0.55	0.40	0.37	0.45	0.35	0.25	0.51
Arithmetic			1	0.48	0.40	0.30	0.26	0.41	0.23	0.29	0.51
Vocabulary				1	0.61	0.38	0.39	0.43	0.33	0.29	0.61
Comprehension					1	0.35	0.34	0.38	0.29	0.23	0.65
Picture Completion						1	0.34	0.47	0.41	0.15	0.48
Picture Arrangement							1	0.41	0.37	0.22	0.38
Block Design								1	0.56	0.30	0.61
Object Assembly									1	0.20	0.58
Coding										1	0.53

[a] Race variable is the point biserial correlation of the dichotomously scored race variable (Zimbabwean Blacks = 1; American Whites = 2) with each of the WISC-R subtests. It is calculated by the formula rpbs = $d/(2)\ (d^2/4) + 1$ where d is the mean standardized difference between Blacks and Whites.

Table 3
Principal factor analysis of correlation matrix for US Whites plus the American White–Zimbabwean Black race variable from Table 2

WISC-R Sub-tests	*g*	Non-*g* factors		
		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
% Total variance	44.073	6.164	4.131	2.228

Africa. The correlation between the sub-tests' *g* loadings and the standardized African–White differences were in the predicted direction but failed to confirm the relationship ($r = 0.36$; ns). This may be due to a restriction of range on African sub-test correlations just as there is in the lower half of the WISC-R IQ distribution of the American standardization sample, where a much larger percentage of the variance is attributable to *g* (Jensen, in press). Regardless, a principal factor analysis of the correlation matrix from the US standardization sample for Whites, along with the effect sizes (point-biserial correlations) of the African–White differences, amply confirms that the African-White difference is primarily on *g*.

Spearman's hypothesis states that the Black–White group difference is more attributable to *g* than to any other psychometric factor. In Jensen's (1987, p. 514, Table 1) analysis of 18 US studies, the average loading of race on *g* was 0.55, much lower than the African loading of 0.88. A factor loading of 0.45 in Jensen (1987) corresponds to a mean Black–White difference of one S.D. This implies that fully 77% of the between-group race variance (African Blacks–American Whites) is attributable to the *g* factor. The loading of 0.88 for the race variable in the present study is much larger than for any of the sub-tests and appears to reflect almost pure *g*. In short, it is strong support for Spearman's hypothesis.

Two blemishes in the study can be noted. First, the comparison was made of Black Zimbabweans to White Americans. This begs the question of whether the same magnitude of difference would be found between Black and White Zimbabweans. A control for nationality would have been ideal. Clinical practice in South Africa has led to the generally accepted view among psychologists that White South Africans do in fact perform less well than White Americans on psychological tests, including the WISC-R (e.g. Nell, 2000). Second, the WISC-R was used and not the WISC-III, a more recent version of the Wechsler Scale. In both cases the study was limited to the data available. These issues may be worthy of further investigation.

In conclusion, the data analysis in this study supports the view that mean sub-Saharan African–White differences on tests of cognitive ability is predominantly a difference in the *g* factor, as Jensen (1998) and Spearman (1927) hypothesized. It is the first such analysis from Zimbabwe, and is in accord with five previous studies from South Africa (Lynn & Owen, 1994; Rushton, 2001, in press; Rushton & Skuy, 2000; Rushton et al., in press). It shows that Spearman's hypothesis applies in southern Africa as it does in the US It appears that the main source of population differences in cognitive performance tests from around the world, is likely the same as that for differences between individuals within each group, namely, Spearman's *g*.

References

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

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

Jensen, A. R. Regularities in Spearman's law of diminishing returns. *Intelligence* (in press).

Jensen, A. R., & Reynolds, C. R. (1982). Race, social class and ability patterns on the WISC-R. *Personality and Individual Differences*, *3*, 423–438.

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.

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

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.

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

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. (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. Jensen effects and African/Colored/Indian/White differences on Raven's Standard Progressive Matrices in South Africa. *Personality and Individual Differences* (in press).

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. Jensen Effects among African, Indian, and White engineering students in South Africa on Raven's Standard Progressive Matrices. *Intelligence* (in press).

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. (1927). *The abilities of man: their nature and measurement*. New York, NY: Macmillan.

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children-Revised*. New York, NY: Psychological Corporation.

Zindi, F. (1994). Differences in psychometric performance. *The Psychologist: Bulletin of the British Psychological Society*, *7*, 549–552.