ELSEVIER

Intelligence 31 (2003) 123–137

INTELLIGENCE

Performance on Raven's Advanced Progressive Matrices by African, East Indian, and White engineering students in South Africa

J. Philippe Rushton[a,*], Mervyn Skuy[b,1], Peter Fridjhon[b]

[a]*Department of Psychology, University of Western Ontario, London, Ontario, Canada N6A 5C2*
[b]*Division of Specialized Education, University of the Witwatersrand, Johannesburg 2050, South Africa*

Received 1 November 2001; received in revised form 19 July 2002; accepted 23 July 2002

Abstract

The hypothesis is tested that the Raven's Advanced Progressive Matrices (APM) has the same construct validity in African university students as it does in non-African university students. Analyses were made of scores from 294 highly select 17–23-year-olds in the Faculties of Engineering and the Built Environment at the University of the Witwatersrand (187 Africans, 40 East Indians, 67 Whites; 70 women, 224 men). Out of a total of 36 problems, the African students solved an average of 22, the East Indian students, 25, and the White students, 29 ($P < .001$), placing them at the 57th, 64th, and 86th percentiles, respectively, and yielding IQ equivalents of 103, 106, and 117 on the 1993 US norms. Four months earlier, they had completed the Standard Progressive Matrices. The two tests correlated .60 or higher for both the Africans and the non-Africans, and both tests predicted final end-of-year grades with mean r's = .30 (P's $< .05$). Items found difficult by one group were difficult for the others; items found easy by one group were easy for the others (mean r's = .90, $P < .001$). The African–East Indian–White differences were "Jensen Effects," being most pronounced on the general factor of intelligence (measured in this instance by items with the highest item-total correlations). Indeed, the *g* loadings showed cross-cultural generality: For example, item-total correlations calculated on the East Indian students predicted the magnitude of the African–White differences. When each of the

* Corresponding author. Tel.: +1-519-661-3685; fax: +1-519-850-2302.
E-mail addresses: Rushton@uwo.ca (J.P. Rushton), 135skuy@mentor.edcm.wits.ac.za (M. Skuy).
[1] Tel.: +27-11-716-5286; fax: +27-11-339-3844.

PII: S0160-2896(02)00140-X

36 *g* loadings and race effects were aggregated into nine four-item "subtests," the magnitude of the Jensen Effect was Spearman's ρ = .52. There were no sex differences.

Keywords: Intelligence; University students; Africa; *g* loadings; Jensen effects

1. Introduction

The average IQ obtained in studies of sub-Saharan Africans is 15–30 points (1–2 S.D.) lower than elsewhere in the world. In their book, *IQ and the Wealth of Nations*, Lynn and Vanhanen (2002) reviewed over two-dozen studies from West, Central, East, and Southern Africa and found they yield an average IQ of around 70. For example, in Nigeria, Fahrmeier (1975) collected data on 375 6–13-year-olds in a study of the effects of schooling on cognitive development. The children's mean score on the Colored Progressive Matrices was 12 out of 36, giving them an IQ equivalent of < 70. In Ghana, Glewwe and Jacoby (1992) reported on a World Bank study that tested a representative sample of 1736 11–20-year-olds from the entire country. All had completed primary school; half were attending "middle-school." Their mean score on the Colored Progressive Matrices was 19 out of 36, which gives an IQ equivalent of < 70. In Zimbabwe, Zindi (1994) gave the Wechsler Intelligence Scale for Children—Revised (WISC-R) and the Standard Progressive Matrices to 204 African 12–14-year-olds, and reported mean IQ scores of 67 on the WISC-R and 72 on the Matrices. In South Africa, Owen (1992) found that 1093 African high school students solved 28 out of 60 problems on the Standard Progressive Matrices, which is around the 10th percentile, or an IQ equivalent of about 80.

Studies published since Lynn and Vanhanen's (2002) review continue to find low scores. In Kenya, Sternberg et al. (2001) administered the Colored Progressive Matrices to 85 12–15-year-olds who scored 23.5 out of 36, an IQ equivalent of about 70. In Tanzania, Sternberg et al. (2002) administered the Wisconsin Card Sorting Task to 358 11–13-year-olds who obtained a mean score on Perseverative Errors of 18.53. Although procedural differences may make the normative comparison problematic, as it stands, this score is equivalent to the fifth percentile on American norms for 12-year-olds (IQ = 75). After receiving training on how to sort attributes, the children's scores went up to 16.5 (lower scores meant fewer errors), but this was still only at the ninth percentile on American norms (IQ < 80). In South Africa, Skuy, Schutte, Fridjhon, and O'Carroll (2001) found mean scores 1–2 S.D. below US norms on a wide variety of tests individually administered to 154 African high school students under optimized conditions. These tests included the Wechsler Intelligence Scale for Children—Revised (WISC-R), the Rey Auditory Verbal Learning Test, the Stroop Color Word Test, the Wisconsin Card Sorting Test, the Bender Gestalt Visual Motor Integration Test, the Rey Osterreith Complex Figure Test, the Trail Making Test, the Spatial Memory Task, and various Drawing Tasks. On the WISC-R, the African students averaged − 1.81 S.D. below American norms (− 1.58 S.D. with the vocabulary subtest excluded).

University students in South Africa also show 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, 1.5 S.D. below US norms (Avenant, 1988; cited by Nell, 2000, pp. 26–28). A study at the University of Venda in South Africa's Northern Province by Grieve and Viljoen (2000) found 30 students in fourth-year law and commerce averaged a score of 37 out of 60 on the Standard Progressive Matrices, equivalent to an IQ equivalent of 78 on US norms. A study at South Africa's University of the North by Zaaiman, van der Flier, and 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, equivalent to an IQ 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 on the basis of a mathematics and science selection test.

At the University of the Witwatersrand in Johannesburg, Rushton, Skuy, and colleagues gave Raven's Standard Progressive Matrices in three separate studies under optimal testing conditions. Rushton and Skuy (2000) found 173 first-year psychology students averaged an IQ equivalent of 84. Skuy et al. (2002) tested another 70 psychology students who averaged an IQ equivalent of 83. After receiving training on how to solve Matrices-type items, their mean scores rose to an IQ equivalent of 96. Rushton, Skuy, and Fridjhon (2002) gave 198 African first-year engineering students the Standard Raven's and found they averaged an IQ of 97. (In contrast, the White university students in these three studies had IQs from 105 to 111; East Indian students had intermediate IQs, from 102 to 106.)

Critics, however, claim that Western-developed IQ tests are not valid for groups as culturally different as sub-Saharan Africans (e.g., Kamin, 1995; Nell, 2000). Thus, Nell (2000, pp. 35–46) adopted a model of "radical environmentalism" and "cultural-relativism" based on Marx, Vygotsky, and Luria (e.g., Luria, 1979) to predict that test items have different meanings for Africans than they do for non-Africans, and that the reason Africans get lower scores is because they are less test-wise: less interested, more anxious, work less efficiently, or give up sooner on items they find difficult than do non-Africans. Nell has gone so far as to argue that standardized tests should not be used in South Africa.

Many of the issues now being discussed in the South African context were previously raised about racial-group differences in the US. For the US, however, many psychometricians have concluded that the tests *are* valid measures of racial differences, at least among people sharing the culture of the authors of the test (Jensen, 1980; Wigdor & Garner, 1982). An 11-person Task Force established by the American Psychological Association to state what was "known" about IQ tests endorsed that assessment. They wrote: "Considered as predictors of future performance, the tests do not seem to be biased against African Americans" (Neisser et al., 1996, p. 93). It is an empirical question how far this conclusion can be universally generalized. It would be useful to have more construct validity data available from other regions of the world, especially since so much new research is being carried out in non-Western countries (e.g., Lynn & Vanhanen, 2002), and controversial findings are emerging (e.g., that the world average IQ is about 90 and that the mean sub-Saharan African IQ is about 70).

Questions of bias fall into two main types: those *internal* to the test and those *external* to it. With regard to internal bias, Nell's (2000) hypothesis of cultural-relativism has been tested and found wanting. Thus, the psychometric studies of Owen (1992) on thousands of high school students, and of Rushton and Skuy (2000, 2002; Skuy et al., 2002) on hundreds of university students, found almost identical item structures in Africans and non-Africans on the Progressive Matrices. Items found difficult by one group were difficult for the others; items found easy by one group were easy for the others (mean r's = .90, P < .001). The item-total score correlations for Africans, East Indians, and Whites were also similar, showing the items measured similar psychometric constructs in all three groups. The only reliable example of bias so far discovered in this extensive literature is the rather obvious internal bias on the Vocabulary components of tests like the Wechsler for groups that do not have English as their first language (e.g., Skuy et al., 2001). Even here, the language factor only accounts for about 0.5 of a standard deviation, out of the overall 2.0 standard deviation difference, between Africans and Whites.

Moreover, six studies in sub-Saharan Africa have supported "Spearman's (1927) hypothesis" that Black–White IQ differences are mainly on *g*, the general factor of intelligence. Lynn and Owen (1994) were the first to find that Africans and Whites differed mainly on the *g* factor in their analysis of data from over 3000 African, East Indian, and White high-school students given 10 subtests of the South African Junior Aptitude Test. Subsequently, Rushton and Skuy (2000) studied 309 South African university students and found that the more the items from the Standard Progressive Matrices measured *g* (estimated by item-total correlations), the more they were related to standardized African–White differences. Rushton (2001) re-analyzed data published on 154 high-school students in South Africa by Skuy et al. (2001) and found African–White differences on 10 subtests of the WISC-R were mainly on *g*. Rushton (2002) analyzed item data from 4000 high-school students in South Africa on the Standard Progressive Matrices published by Owen (1992) and found the four-way African–Colored–East Indian–White mean differences were all on *g*. Rushton et al. (2002) studied 342 South African engineering students and again found the more *g* loaded items from the Standard Progressive Matrices (again estimated from item-total correlations) better predicted African–East Indian–White differences. Finally, Rushton and Jensen (in press) compared data from 204 African 12–14-year-olds from Zimbabwe published by Zindi (1994) with the US normative sample for Whites on the WISC-R and found 77% of the between-group race variance was attributable to a single source, namely *g*. All the above studies showed that the items measured the same underlying construct in both the African and the non-African groups.

With regard to claims of external bias, a review by Kendall, Verster, and Von Mollendorf (1988) showed that test scores for Africans have about equal predictive validity as they do for non-Africans (e.g., 0.20–0.50 for school grades in students and for job performance in employees). Kendall et al. also showed that many of the factors that influence scores in Africans are the same as those for Whites (e.g., coming from an urban versus a rural environment, being a science rather than an arts student, having had practice on the tests, and the well-documented curvilinear relationship with age). Similarly, Sternberg et al.'s (2001) study of Kenyan 12–15-year-olds found that IQ scores predicted school grades with a mean r = .40 (P < .001; after controlling for age and socioeconomic status, r = .28, P < .01).

The present study examines further whether "universalist" assumptions are valid in the South African context. The construct validity of IQ scores is examined in a highly select group of African, East Indian, and White university students on the more difficult Raven's Advanced Progressive Matrices (APM). *External* validity is assessed by examining whether scores on the APM predict final end-of-year examination marks taken 4 months later, as well as scores gathered on the Standard Progressive Matrices 4 months earlier. *Internal* validity is assessed by examining whether the item structures and *g* loading patterns are the same for the different racial groups, and whether African–White group differences are most pronounced on the *g* factor as hypothesized by Spearman (1927) (see Jensen, 1998), and observed in previous studies. One of the hypotheses raised by Rushton et al. (2002) is also tested, namely that a higher performing group of African students might be identified, yielding evidence of a bimodal distribution in African university students. The APM is an ideal test for these purposes (see more below) because it was developed to spread out the distribution of scores in the top 25% of the population (Raven, 2000; Raven, Raven, & Court, 1998).

2. Method

2.1. Overview

The primary purpose was to examine performance on the APM in an African sample that would be expected to score very much above the general South African population mean. In the US, engineers are among the very highest scoring groups on tests such as the Scholastic Aptitude Test (SAT) and the Graduate Record Examination (GRE). For example, the mean verbal + quantitative + analytic scores of engineering students on the GRE is about 1800, whereas psychology and education students average about 1500, a difference of about one standard deviation (Educational Testing Service, 1998). In turn, students entering university to study psychology score about one standard deviation above the general population. Thus, engineers may be up to about two standard deviations above the general population, at least in the US and most developed nations.

First-year students from the Faculties of Engineering and the Built Environment at the University of the Witwatersrand are among the highest academically achieving African students. They had been tested on Raven's Standard Progressive Matrices in April 2001 (Rushton et al., 2002) and were invited to take part in a follow-up study in August 2001. Participants were paid 50 rand (at that time about US$8) each. Final end-of-year examination marks were taken in December 2001. (In the Southern Hemisphere, December marks the completion of the academic year and the beginning of the long summer break.)

2.2. Subjects

An initial pool of 342 subjects who took the APM was reduced to 294 17–23-year-olds by eliminating small or ambiguous categories. Excluded were those who had not taken part in the earlier testing ($n=31$), or who, in the earlier testing, had self-identified as "Colored" ($n=5$), "Other" ($n=7$), listed their age as over 23 ($n=2$), or failed to give biographical data ($n=3$). The analyses were conducted on 187 Africans (144 men, 43 women), 40 East Indians (24 men, 16 women), and 67 Whites (56 men, 11 women). (Forty-eight of those who had taken the Standard Progressive Matrices did not turn up for the APM: 11/198 = 5.5% of the Africans, 18/58 = 31% of the East Indians, and 19/86 = 22% of the Whites).

2.3. Test instrument

The Raven's APM is the most difficult of the Raven's Matrices tests. Regardless of version, the Raven's tests (the Colored Progressive Matrices, the Standard Progressive Matrices, and the APM) are the best known, most researched, and most widely used of all culture-reduced tests (Raven, 2000). The Raven's APM was designed to differentiate between people of "superior intellectual ability," such as "students for advanced scientific or technical studies" (Raven et al., 1998, pp. 1–2). The Advanced Matrices test shows reliability and validity across a range of populations, including African Americans and other non-Whites. The total score provides a good measure of *g*, the general factor of intelligence, at least within the U.S. (Jensen, 1980; Raven et al., 1998).

The APM consists of two components, published in separate booklets. Set 1 consists of a short set of 12 diagrammatic puzzles, each with a missing part that the test taker attempts to identify from several options. Set 1 calls for all the intellectual processes sampled on the full test. It is normally used for practice, to familiarize people with the test, and to allay anxiety. In this study, Set 1 with the correct answers was given to the students to take home 1 week before the testing session so they could practice and test themselves. Four months earlier, these students had also been given the Raven's Standard Progressive Matrices (Rushton et al., 2002).

Set II consists of 36 puzzles that are identical in presentation with those in Set 1. However, they increase in difficulty more steadily and become considerably more complex. As the items are arranged in inverse order to the frequency with which they are solved, the validity of the total score does not depend (as it does on the Standard version) upon everyone attempting all the items in the Set. To ensure sustained interest and freedom from fatigue, each problem is boldly presented, accurately drawn, and, as far as possible, pleasing to look at.

2.4. Procedure

The Raven's APM was administered without any time limits (up to 1 h), but was typically completed within 30 min. Professor Skuy, one of the authors of this study, and his colleagues, carried out the testing in a classroom. All students appeared well motivated. The instructions requested students to wait quietly at their desks if they finished before 30 min. After 30 min, however, they could come to the front of the room, hand in their answer sheets and test booklets, and receive payment.

3. Results

3.1. Means, S.D., and construct validities

All calculations are based on raw scores, with each of the 36 items scored as 0 (*incorrect*) or 1 (*correct*). Internal consistencies based on Cronbach's α were .85 for the sample as a

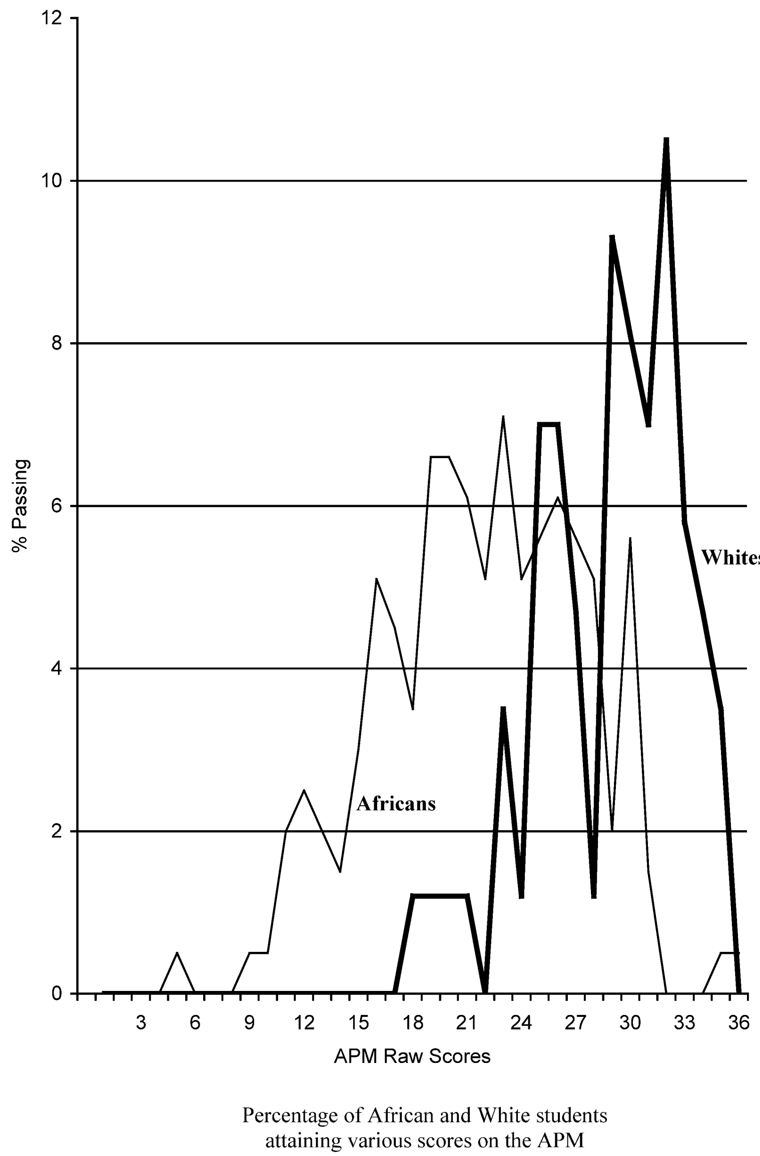

Fig. 1. Percentage of African and White 17–23-year-old first-year engineering students attaining various scores on the Raven's Advanced Progressive Matrices Test.

whole ($n = 294$), .82 for the Africans ($n = 187$), .79 for the East Indians ($n = 40$), and .75 for the Whites ($n = 67$). The 48 no-shows for the second testing did not have significantly different mean scores on the Standard Progressive Matrices than those who showed up for both sessions.

The African, East Indian, and White mean scores were, in order, 22, 25, and 29 out of 36 (S.D. = 5.6, 4.9, 4.0; ranges = 5–36, 12–34, 18–35). Men averaged the same scores as women (unweighted means = 23.9, 22.9; S.D. = 5.5, 6.0; ranges = 5–36, 11–33). Analysis of variance (ANOVA) with race and sex as factors showed a significant main effect only for race, with no effect for sex either as a main effect or an interaction, $F(2,294) = 28.51$, $P < .001$; $F(1,294) < 1.00$; and $F(2,294) < 1.00$. For the total score, the African–White difference was 1.54 S.D. (based on total S.D. of 5.92). Using the 1993 US norms for 18–22-year-olds, the Africans fall at the 57th percentile, the East Indians at the 64th percentile, and the Whites at the 86th percentile, which yield IQ equivalents of 103, 106, and 117, respectively (Raven et al., 1998). Post-hoc tests showed that both the African–East Indian and East Indian–White means were significantly different from each other ($P < .05$).

Fig. 1 shows the percentage of Africans and of Whites who attained various raw scores (with the intermediately scoring East Indian group being excluded for clarity). The Africans show an almost normal distribution on this test, although the longer tail of low scores is visible; the ceiling effect remains for the White group. There is no evidence of a bimodal distribution among the Africans with a higher scoring group equal to the Whites, as conjectured might be the case by Rushton et al. (in press).

4. Item analyses

Table 1 shows the proportion of each of the samples that selected the correct answer on each of the 36 items. For all groups, the test items increased in difficulty as the test progressed (mean $r = .90$, $P < .001$). The item difficulties, measured by the proportion getting the correct

Table 1
Proportion of sample selecting the correct answer on items of the Advanced Progressive Matrices by race

Item	African	Indian	White	Item	African	Indian	White	Item	African	Indian	White
1	.94	1.00	1.00	13	.68	.85	.93	25	.49	.50	.75
2	.89	.97	.99	14	.74	.85	.99	26	.45	.35	.60
3	.95	.97	.99	15	.81	.80	.90	27	.35	.40	.69
4	.90	.93	.94	16	.75	.75	.99	28	.27	.20	.51
5	.90	.87	.99	17	.72	.87	.88	29	.24	.15	.43
6	.92	1.00	.99	18	.53	.53	.75	30	.27	.38	.69
7	.90	.95	.99	19	.69	.62	.81	31	.28	.42	.52
8	.76	.90	.96	20	.65	.70	.78	32	.19	.28	.45
9	.90	1.00	1.00	21	.60	.75	.93	33	.32	.38	.54
10	.71	.88	.96	22	.58	.60	.94	34	.30	.30	.57
11	.84	.95	.97	23	.59	.70	.90	35	.32	.47	.63
12	.80	.92	.99	24	.42	.45	.75	36	.09	.18	.28

Table 2
Item-total correlations for the Advanced Progressive Matrices by race

Item	African	Indian	White	Item	African	Indian	White	Item	African	Indian	White
1	.09	–	–	13	.24	.11	.23	25	.50	.30	.38
2	.29	.09	.31	14	.53	.60	.18	26	.37	.46	.30
3	.29	.09	.09	15	.38	.21	.32	27	.45	.29	.39
4	.34	−.01	.32	16	.54	.56	.24	28	.28	.31	.39
5	.15	.10	.09	17	.34	.41	.14	29	.24	−.30	.46
6	.18	–	.18	18	.36	.54	.38	30	.42	.35	.59
7	.47	.46	.24	19	.34	.54	.30	31	.22	.43	.44
8	.32	.23	.48	20	.29	.40	.13	32	.30	.49	.52
9	.46	–	–	21	.56	.45	.45	33	.28	.46	.48
10	.55	.58	.36	22	.46	.53	.32	34	.38	.47	.50
11	.47	.35	.13	23	.56	.48	.19	35	.46	.56	.53
12	.40	.34	.34	24	.38	.50	.48	36	.24	.19	.24

Hyphen indicates that correlation could not be computed because of lack of variance on item (see Table 1).

answer (Table 1), were similar for Africans, East Indians, and Whites (r's > .90, P's < .001). This suggests that the test measures the same construct in all three groups. However, even on this Advanced test, 55% of the items were too easy for these university students (as indicated by pass rates of 71% or more; 35% of the items had an optimal pass rate of between 30% and 70%, which provides maximum discriminatory power; while 10% of the items were too difficult, with a pass rate between 0% and 29%). Only one item (item 36 for the Africans) was "extremely difficult" (P < .10) for any of the groups.

Another index for comparing items across groups is the item-total correlation (r_{it}) calculated using both the point-biserial (r_{pb}) (Table 2) and the biserial correlations (r_b) (Table 3) of each item's pass or fail status (0 or 1) with the total score on the test. These indicate the extent to which a particular item measures the same construct that is measured by the test as a whole, as well as how well the item discriminates among the testees within each group. Since

Table 3
Biserial item-total correlations for the Advanced Progressive Matrices by race

Item	African	Indian	White	Item	African	Indian	White	Item	African	Indian	White
1	.18	–	–	13	.32	.16	.43	25	.61	.37	.53
2	.47	.25	.93	14	.69	.92	.53	26	.49	.58	.39
3	.60	.25	.27	15	.55	.33	.51	27	.60	.32	.49
4	.60	−.04	.64	16	.73	.78	.77	28	.39	.42	.45
5	.23	.16	.27	17	.47	.63	.24	29	.34	−.48	.56
6	.32	–	.53	18	.45	.69	.53	30	.54	.41	.76
7	.80	.98	.77	19	.43	.68	.41	31	.31	.55	.56
8	.42	.37	1.00	20	.40	.53	.16	32	.44	.66	.67
9	.80	–	–	21	.68	.60	.83	33	.39	.59	.50
10	.74	.89	.76	22	.57	.68	.64	34	.47	.62	.62
11	.72	.71	.31	23	.73	.62	.32	35	.59	.69	.66
12	.56	.62	1.03	24	.48	.64	.63	36	.42	.29	.27

Hyphen indicates that correlation could not be computed because of lack of variance on item (see Table 1).

the total score on the Raven's is a very good measure of *g*, the general factor of intelligence (Jensen, 1980, pp. 645–648), the item-total correlation is also an estimate of each item's *g* loading. (The r_b correlation is used here to calculate item-total correlations in addition to the more usual r_{pb} because its formula contains within it a correction for item difficulty in the form of the ordinate, *y*, of the normal curve at the % pass/% fail cut-point. Since r_{pb} lacks this correction, it is possible that r_{pb} confounds *g* loading with item-difficulty level, thus, making the apparent Jensen Effects an artifact of item difficulty rather than a true reflection of *g*.)

5. Differences in *g*

To test whether African–East Indian–White differences are more pronounced on the more *g* loaded items, we followed the same procedure as in our previous studies (Rushton & Skuy, 2000; Rushton et al., 2002), and correlated the item-total correlations from Tables 2 and 3 (the estimates of *g*), with the standardized differences between the ethnic groups in proportion passing each item from Table 1 (the estimate of the race effect size). The results are set out in Table 4 using, in turn, the African, East Indian, and White point-biserial (r_{pb}) and biserial (r_b) item-total correlations. The 36 Pearson *r*'s and Spearman ρ's ranged from .08 to .64, with a mean of .31 and a median of .26, with 22 of the 36 correlations being at least marginally significant ($P < .10$). The African–White differences were on *g* regardless of whether *g* was calculated from African, East Indian, or White point-biserial (r_{pb}) correlations (mean $r = .48$, mean $\rho = .49$, mean P's $< .01$) or by African, East Indian, or White biserial (r_b) correlations (mean $r = .38$, mean $\rho = .39$, mean P's $< .05$). (Note that it would have been incorrect to use the item-total correlations from the *combined* samples because these would reflect the *between*-groups variance in addition to the *within*-groups variance and so inflate the effect.)

Table 4
Point-biserial (r_{pb}) and biserial (r_b) item-total correlations (*g* loadings) and differences in standardized African–Indian–White item pass rates on the Raven's Progressive Matrices

	Item-total correlations					
	African ($n=36$)		Indian ($n=33$)		White ($n=34$)	
	r_{pb}	r_b	r_{pb}	r_b	r_{pb}	r_b
Pearson correlations						
Indian–White	.509**	.265†	.293 *	.175	.198	.209
African–Indian	.212	.174	.254†	.278†	.196	.251†
African–White	.637**	.379*	.469**	.377*	.334*	.385*
Spearman correlations						
Indian–White	.538**	.316*	.324*	.195	.212	.191
African–Indian	.165	.081	.162	.181	.213	.257†
African–White	.640**	.371*	.499**	.424**	.339*	.382*

* $P < .05$.
** $P < .01$.
† $P < .10$.

Low reliabilities found for data at the item level likely depressed these Jensen Effects (as *g* factor effects are called). Since in the Rushton et al. (2002) study, aggregating both the item-total correlations (the measure of *g*) and the race effects enhanced the magnitude of the Jensen Effect, aggregation was employed here to offset the depression effect of low item reliability. Scores at the item level were averaged four at a time to make nine "subtests." This procedure boosted the size of the Jensen Effects on the Black–White difference regardless of whether calculated from African, Indian, or White point-biserial (r_{pb}) correlations (mean $r = .55$, mean $\rho = .52$, mean P's $< .10$) or by African, Indian, or White biserial (r_b) correlations (mean $r = .55$, mean $\rho = .52$, mean P's $< .10$).

6. External validities

External validities were established. For both the African and the non-African students, scores on the Advanced Matrices predicted scores on the Standard Matrices measured 4 months earlier, with $r = .60$ for the 187 Africans and $r = .70$ for the 107 non-Africans (.66 for the 40 East Indians and .58 for the 67 Whites). Even more importantly, the Advanced Matrices also predicted final end-of-year exam marks measured 4 months later, with $r = .34$

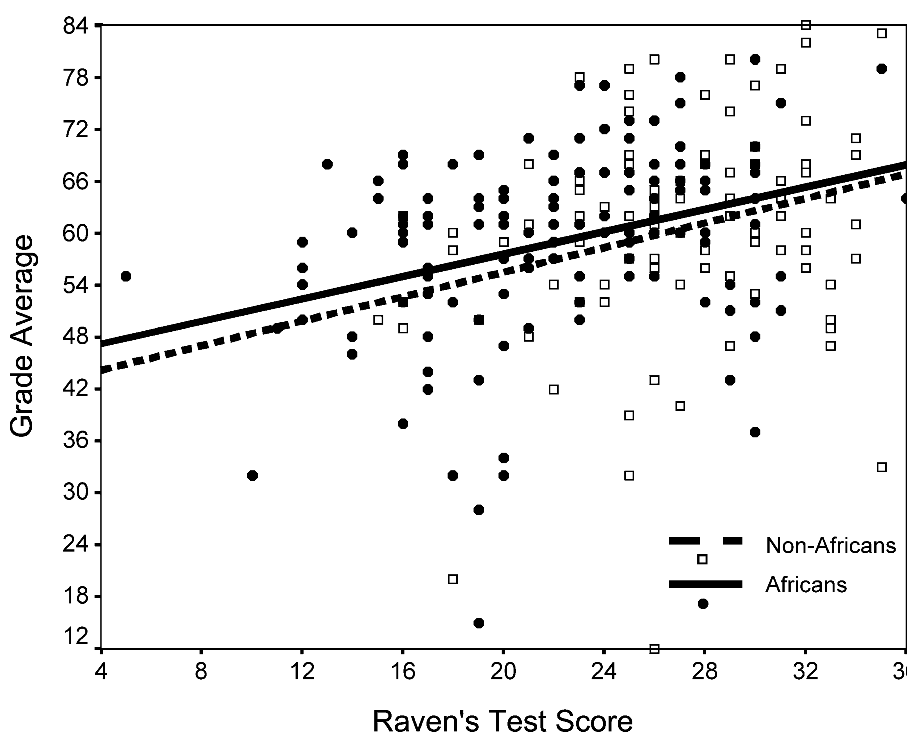

Fig. 2. Regression of Raven's scores on university grades for Africans and non-Africans.

for the Africans and r = .28 for the non-Africans (.18 for the East Indians and .08 for the Whites, who had greater restriction of range). Fig. 2 shows the regression of exam marks on test scores for the African and the non-African university students (with the scores from the smaller samples of East Indian and White students combined for clarity). Because the African and non-African regression lines run almost completely parallel over the entire range of scores, these results indicate that the tests are likely not biased against Africans.

7. Discussion

The African engineering students at the University of the Witwatersrand are the highest scoring African sample measured to date, solving an average of 22 out of the 36 problems on Set II of the APM. The East Indian and White students solved 25 and 29 out of 36, respectively. By the 1993 US norms, the African students were at the 57th percentile, the Indians at the 64th, and the Whites at the 86th, yielding IQ equivalents of 103, 106, and 117, respectively.

As expected, these scores are all higher than those obtained 4 months earlier on the Standard Progressive Matrices, with the Africans gaining the equivalent of six IQ points and the non-Africans gaining five. These gains could be due to a combination of the removal of the ceiling effect, and to the practice effect of having taken the Standard Progressive Matrices 4 months earlier and Set I of the APM 1 week earlier. The African–White difference amounted to 1.54 S.D. (based on the total sample S.D. of 5.92). There was no evidence of a bimodal distribution among the African students or of a higher scoring African group equal to Whites, as conjectured might be the case by Rushton et al. (2002).

The construct validity of African IQ scores was confirmed in this study. The inter-item matrices were the same for Africans as for the non-Africans, with similar α coefficients, item difficulty levels, and g factor loadings. External validities were also established, with the African students' APM scores predicting their scores on the Standard Progressive Matrices taken 4 months earlier (r = .60) and their final end-of-year examination marks taken 4 months later (r = .34). These results thus join those from earlier studies showing that mental ability test scores in Africans do have predictive power (Kendall et al., 1988; Sternberg et al., 2001).

The data from the present study confirm that African university students range in IQ score from 77 to 103. As university students in Africa likely score 1–2 S.D. above the general average, their scores are consistent with the results of previous studies which found the mean IQ score of the African general population to be around 70 (Lynn & Vanhanen, 2002). The present study also confirms the findings that African–White differences reflect g rather than culturally specific ways of thinking. The g effect is robust and implies that African–White differences are due to factors similar to those in other countries, such as the US and the Netherlands (Jensen, 1998; te Nijenhuis & van der Flier, 2001).

Questions can be raised, however, about these conclusions. One reviewer suggested that persons with such low IQs could not complete law and medical school and then practice these professions and suggested, therefore, that the low scores called into question the predictive validity of the tests. Two explanations for this conundrum are possible. The first is that the low African mean IQ score *does* represent the probable level of cognitive performance for the population and that, indeed, commensurate work performance is expected (see Lynn & Vanhanen, 2002). The second is that although individual differences in IQ score *within* a population are predictive of individual differences in various criteria within that population (as in Fig. 2), differences *between* populations are due to poverty and cognitive privation and that high motivation would be able to outweigh scores in determining performance. According to this view, Euro–American test norms are not valid for Africans. Future research using some of the real-life criteria identified by Gordon (1997) and Gottfredson (1997) may shed light on the issue.

Another reviewer questioned whether the *g* factor extracted from the homogeneous Raven items found on the Advanced Matrices is representative of the psychometric *g* that would be found in a factor analysis of a wider variety of mental tests. Evidence on this issue was provided by Vernon (1983), who found that when the Advanced Matrices was included among the 11 subtests of the WAIS and administered to 100 university undergraduates at the University of California at Berkeley (mean IQ = about 120), the Matrices had the largest *g* loading of any of the various subtests, even in this range-restricted sample. Thus, the Advanced Matrices is a good *g* indicator, even in college-selected samples.

The fact that African–East Indian–White differences are on the *g* factor certainly does not preclude the use of intervention strategies to ameliorate cognitive deficits and thereby narrow group differences. The main suggestions proposed for raising cognitive performance in disadvantaged groups include improving nutrition, welfare, hygiene, and schooling, and some of these have been discussed by Lynn and Vanhanen (2002) and Raven (2000). Importantly, Raven also reviewed some of the experimental research showing that students who are encouraged to engage in complex cognitive activity increase their level of self-direction, understanding, and competence. Likewise, Skuy and colleagues used Feuerstein's (1980) Mediated Learning Experience to boost Raven scores in South Africa. Skuy and Shmukler (1987) applied the Mediated Learning Experience to raise the performance of non-White high school students, and Skuy, Hoffenberg, Visser, and Fridjhon (1990) found generalized improvements for individuals who possessed what they termed a "facilitative temperament." In an intervention study with first-year psychology students at the University of the Witwatersrand, Skuy et al. (2002) similarly raised test scores in both Africans and non-Africans who were administered the Raven's on two separate occasions. After intervention training, both the African and the non-African experimental groups improved over baseline compared to the control groups, with significantly greater improvement for the African experimental group.

From a *g* perspective, the question is whether these and other intervention techniques boost *g* or more specific skills (Raven, 2000). In the Netherlands, te Nijenhuis and van der Flier (2001) reported that immigrant/non-immigrant group differences on the *g* factor had narrowed by the second generation of immigrants. In the US, there is some evidence that the Black–White difference may also be narrowing (Jencks & Phillips, 1998; Neisser, 1998), although it is not known whether this is on the *g* factor. Regardless, from a pragmatic perspective, the present paper shows that African–White differences are not attributable to idiosyncratic cultural peculiarities in this or that test. Rather, they reflect the general factor that all ability tests measure.

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.

Educational Testing Service (1998). *Graduate Record Examinations: guide to the use of scores*. Princeton, NJ: Educational Testing Service.

Fahrmeier, E. D. (1975). The effect of school attendance on intellectual development in Northern Nigeria. *Child Development*, *46*, 281–285.

Feuerstein, R. (1980). *The dynamic assessment of retarded performers*. Glenview, IL: Scott, Foresman.

Glewwe, P., & Jacoby, H. (1992). *Estimating the determinants of cognitive achievement in low-income countries: the case of Ghana*. Washington, DC: World Bank.

Gordon, R. A. (1997). Everyday life as an intelligence test: effects of intelligence and intelligence context. *Intelligence*, *24*, 203–320.

Gottfredson, L. S. (1997). Why *g* matters: the complexity of everyday life. *Intelligence*, *24*, 79–132.

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.

Jencks, C., & Phillips, M. (Eds.) (1998). *The Black–White test score gap*. Washington, DC: Brookings Institution Press.

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

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

Kamin, L. (1995). Lies, damned lies, and statistics. In R. Jacoby, & N. Glauberman (Eds.), *The bell curve: history, documents, opinions* (pp. 81–105). New York, NY: Random House.

Kendall, I. M., Verster, M. A., & Von Mollendorf, J. W. (1988). Test performance of blacks in Southern Africa. In S. H. Irvine, & J. W. Berry (Eds.), *Human abilities in cultural context* (pp. 299–339). Cambridge, UK: Cambridge Univ. Press.

Luria, A. R. (1979). *The making of mind*. Cambridge, MA: Harvard Univ. Press.

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.

Neisser, U. (Ed.) (1998). *The rising curve: long term gains in IQ and related measures*. Washington, DC: American Psychological Association.

Neisser, U., Boodoo, G., Bouchard Jr., T. J., 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*. London: Erlbaum.

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

Raven, J. (2000). The Raven's Progressive Matrices: change and stability over culture and time. *Cognitive Psychology*, *41*, 1–48.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's Advanced Progressive Matrices (1998 edition)*. Oxford, England: Oxford Psychologists Press.

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., & 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., & 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., & Fridjhon, P. (2002). Jensen Effects among African, Indian, and White engineering students in South Africa on Raven's Standard Progressive Matrices. *Intelligence*, 30, 409–423.

Skuy, M., Gewer, A., Osrin, Y., Khunou, D., Fridjhon, P., & Rushton, J. P. (2002). Effects of mediated learning experience on Raven's Matrices scores of African and non-African university students in South Africa. *Intelligence*, *30*, 221–232.

Skuy, M., Hoffenberg, S., Visser, L., & Fridjhon, P. (1990). Temperament and cognitive modifiability of academically superior black adolescents in South Africa. *International Journal of Disability, Development, and Education*, *37*, 29–44.

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.

Skuy, M., & Shmukler, D. (1987). Effectiveness of the learning potential assessment device with Indian and Colored adolescents in South Africa. *International Journal of Special Education*, *12*, 131–149.

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

Sternberg, R. J., Grigorenko, E. L., Ngorosho, D., Tantufuye, E., Mbise, A., Nokes, C., Jukes, M., & Bundy, D. A. (2002). Assessing intellectual potential in rural Tanzanian school children. *Intelligence*, *30*, 141–162.

Sternberg, R. J., Nokes, C., Geissler, P. W., Prince, R., Okatcha, F., Bundy, D. A., & Grigorenko, E. L. (2001). The relationship between academic and practical intelligence: a case study in Kenya. *Intelligence*, *29*, 401–418.

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.

Vernon, P. A. (1983). Speed of information-processing and general intelligence. *Intelligence*, *7*, 53–70.

Wigdor, A. K., & Garner, W. R. (Eds.) (1982). *Ability testing: uses, consequences, and controversies: Part 1. Report of the Committee, Part 2. Documentation section*. Washington, DC: National Academy Press.

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. *Psychologist*, *7*, 549–552.