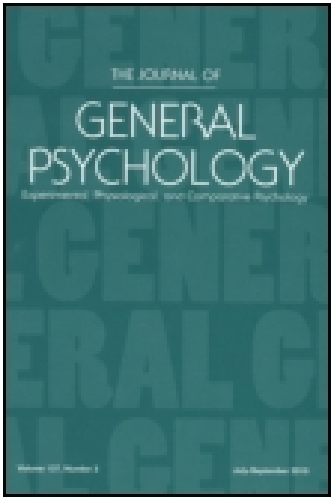
This article was downloaded by: [New York University]
On: 06 December 2014, At: 12:47
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



The Journal of General Psychology

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/vgen20

Spearman's Hypothesis and Test Score Differences Between Whites, Indians, and Blacks in South Africa

Richard Lynn [a] & Kenneth Owen [b]

[a] Psychology Department , University of Ulster at Coltraine

[b] Human Sciences Research Council , Pretoria, South Africa

Published online: 06 Jul 2010.

To cite this article: Richard Lynn & Kenneth Owen (1994) Spearman's Hypothesis and Test Score Differences Between Whites, Indians, and Blacks in South Africa, The Journal of General Psychology, 121:1, 27-36, DOI: 10.1080/00221309.1994.9711170

To link to this article: http://dx.doi.org/10.1080/00221309.1994.9711170

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or

indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

Downloaded by [New York University] at 12:47 06 December 2014

The Journal of General Psychology, 121(1), 27–36

Spearman's Hypothesis and Test Score Differences Between Whites, Indians, and Blacks in South Africa

RICHARD LYNN
Psychology Department
University of Ulster at Coltraine

KENNETH OWEN
Human Sciences Research Council
Pretoria, South Africa

Downloaded by [New York University] at 12:47 06 December 2014

ABSTRACT. Numerous studies in the United States have shown that mean test scores between Blacks and Whites differ by about one standard deviation. It has further been noted that the magnitudes of these differences vary on different tests. This variation can be explained by Spearman's hypothesis, which states that Black–White differences on a set of cognitive tests are positively associated with the tests' *g* loadings (the general intellectual ability). The present study, conducted among Black, Indian, and White secondary students in South Africa, showed mean Black–White differences of two standard deviations, indicating that the American results of one standard deviation are not universally correct. With regard to Spearman's hypothesis, it was found that, although the mean White–Indian differences were about one standard deviation, these differences did not support the hypothesis. Results pertaining to the Black–White differences were ambiguous; the correlation of .62 ($p < .05$) between the Black *g* and the Black–White differences strongly supported the hypothesis. A nonsignificant correlation of .23 was obtained between the White *g* and the Black–White differences. Possible reasons for this finding are discussed.

IT IS WIDELY ACCEPTED THAT in the United States the average mean scores on intelligence tests obtained by Blacks are approximately 15 IQ points lower than those obtained by Whites (Jensen, 1969; Loehlin, Lindzey & Spuhler, 1975; Osborne & McGurk, 1982). In addition, it has been known for a number of years that Black–White differences in intelligence in the United States are more pronounced on tests of some abilities than on others. In particular, the differences

Address correspondence to Richard Lynn, Psychology Department, University of Ulster, Coleraine County Londonberry BT52 1SA, Northern Ireland.

Downloaded by [New York University] at 12:47 06 December 2014

have generally been relatively small on tests of rote learning and immediate memory and greater on tests requiring problem solving and more complex mental operations (Jensen, 1985). Spearman (1927) first noted this Black–White difference and suggested that the magnitude of the difference is positively related to the degree to which tests measure general intellectual ability (g), that is, the more highly correlated a test is with *g*, the greater the Black–White difference. Jensen (1985) designated this proposition *Spearman's hypothesis* and assembled 11 studies on which the hypothesis could be tested. He found that the overall correlation between the tests' *g* loadings and the magnitude of the Black–White difference is +0.59, a statistically highly significant correlation supporting the hypothesis.

All the studies on which Jensen based his analysis were carried out in the United States; therefore, it has not been demonstrated that Spearman's hypothesis holds true for other Black–White populations. If not, the phenomenon is of limited interest and an explanation could be sought in local conditions in the United States. On the other hand, if the hypothesis holds elsewhere, the phenomenon becomes one of more universal validity and interest.

Our aim in the present study was threefold. First, we sought to report mean test score differences for a number of abilities between Black and White South Africans and to compare these score differences with those found in the United States. Second, we wanted to ascertain whether Spearman's hypothesis that Black–White differences are principally differences in *g* holds true for South Africa. Third, we wanted to report intelligence test means for Indians in South Africa, as both a matter of general interest and to examine whether there are Indian–White differences in Spearman's *g*.

Method

Samples

The sample consisted of adolescents aged 15–16 in South African "standard 7" classes in secondary schools. The mean numbers and ages of the subjects were as follows: Whites, $n = 1{,}056$, mean age 15.0 years ($SD = 0.86$); Blacks, $n = 1{,}093$, mean age = 16.5 years ($SD = 1.67$); Indians, $n = 1{,}063$, mean age = 15.0 years ($SD = 0.99$). There were approximately equal numbers of boys and girls in each group. The White subjects were drawn from 20 schools in the Pretoria–Witwatersrand–Vereeniging (PWV) area and 10 schools in the Cape Peninsula. The Indian subjects were drawn from 30 schools selected at random from the list of high schools in and around Durban. The Black sample came from three schools in the Pretoria–Witwatersrand–Vereeniging area and from 25 schools selected as representative schools in Black areas in KwaZulu adjacent to urban centers in Natal. We administered the tests in August and September 1985 in the White and Indian schools, in November 1985 in the three Black schools in the PWV area, and in June 1986 in the 25 schools in KwaZulu. Testing was conducted by school

psychologists from the various groups. Owen (1989) gives further details of the samples.

We used the South African Junior Aptitude Tests (JAT) (Verwey & Wolmarans, 1983); a multiability test constructed in South Africa and standardized on White pupils in Standards 5 to 8. The JAT is similar to the American DAT and consists of 10 tests of primary abilities.

1. Classification: a nonverbal reasoning test based on pictures of objects with certain characteristics in common
2. Reasoning: a test consisting of verbal and numerical reasoning problems
3. Number: a test containing arithmetic problems of addition, subtraction, multiplication, division, and percentages
4. Synonyms: a verbal comprehension test
5. Comparisons: a perceptual speed test consisting of matching groups of letters and numbers with a standard group
6. Spatial 2-D: a two-dimensional spatial-comprehension test involving combining shapes to form a square
7. Spatial 3-D: a test of three-dimensional spatial ability
8. Memory 1 (Paragraph): a test of memory for meaning that involves reading paragraphs and later answering questions on the content
9. Memory 2 (Symbols): an associative memory test that requires remembering arbitrary associations of pairs of words and symbols
10. Mechanical Insight: a mechanical ability test consisting of questions concerning drawings of mechanical apparatus

Further details regarding the tests' reliability, validity, and so on, are given in Owen's monograph and the test manual.

Results

It was evident (see Table 1) that considerable test score differences existed among the three groups (Hotelling's T^2 and post hoc t tests revealed that all White–Indian and White–Black differences were statistically significant, $p < .0001$). The actual size of these differences can be better appreciated when they are expressed as standardized differences (in terms of each test's standard deviation for Whites). These standardized differences also were corrected for attenuation by dividing each difference by the square root of the particular test's reliability coefficient (KR-20) for Whites.

The mean of the standardized Black–White differences on the 10 tests was 2.1 *SD* units, whereas the mean of the White–Indian differences was 1.0 *SD* unit. The White–Black and White–Indian differences were relatively small for the associative memory test (Test 9) and number manipulation test (Test 3), as compared with the other tests, which place more emphasis on problem solving. It was further evident that the nonverbal reasoning test (Test 1) showed the biggest

Downloaded by [New York University] at 12:47 06 December 2014

Downloaded by [New York University] at 12:47 06 December 2014

TABLE 1
Means, Standard Deviations (*SD*s), and Test Reliabilities (KR-20) of the JAT for White, Indian and Black Pupils, Standardized White–Indian and White–Black Test Differences, and Differences Corrected for Attenuation

Tests	Maximum score	White (n = 1,056)			Indian (n = 1,063)			Black (n = 1,093)			Standardized White–Indian differences*	Standarized White–Black differences**	Corrected White–Iindian differences***	Corrected White–Black differences****
		Mean	SD_1	KR-20	Mean	SD_2	KR-20	Mean	SD_3	KR-20				
JAT1: Classification	30	23.1	3.5	0.63	19.1	4.1	0.65	13.4	4.3	0.64	1.14	2.77	1.44	3.51
JAT2: Reasoning	30	18.3	4.8	0.79	13.7	4.7	0.78	6.5	2.6	0.33	0.96	2.46	1.08	2.76
JAT3: Number Ability	30	16.7	4.9	0.80	15.6	4.8	0.80	9.0	3.6	0.66	0.22	1.57	0.25	1.76
JAT4: Synonyms	30	19.9	4.9	0.78	14.5	5.6	0.81	6.1	2.7	0.29	1.10	2.82	1.25	3.20
JAT5: Comparison	30	24.6	3.9	0.79[1]	20.6	5.7	0.77[1]	16.5	5.5	0.60[1]	1.03	2.08	1.16	2.34
JAT6: 2-D	40	27.8	7.2	0.87	17.4	6.9	0.87	13.6	5.9	0.83	1.44	1.97	1.55	2.12
JAT7: 3-D	30	22.0	5.5	0.85	14.5	6.4	0.87	10.3	4.9	0.80	1.36	2.13	1.48	2.32
JAT8: Memory (Paragraph)	25	17.4	1.6	0.80	13.6	4.3	0.73	8.2	3.8	0.66	0.83	2.0	0.93	2.25
JAT9: Memory (Symbols)	30	22.7	5.5	0.87	19.5	5.6	0.84	16.7	5.9	0.84	0.58	1.09	0.62	1.17
JAT10: Mechanical Insight	42	19.7	5.5	0.75	12.2	4.0	0.60	7.9	3.0	0.32	1.36	2.15	1.56	2.47

*$(\bar{X}1 - \bar{X}2)/S1$. **$(\bar{X}1 - \bar{X}3)/S1$. ***$[(\bar{X}1 - \bar{X}2)/S1]/\sqrt{\text{White KR-20}}$. ****$[(\bar{X}1 - \bar{X}3)/S1]/\sqrt{\text{White KR-20}}$.

1) Mean of the other 9 KR-20 values

Black–White difference of all. This finding stresses the point made by Irvine (1969) that items with figural content do not necessarily lead to a reduction in cultural bias. (Whether the difference in the case of the present study was due to cultural bias or to lack of ability is not easily decided.)

The appreciable differences between Black and White scores on Test 2 (mainly verbal reasoning), Test 4 (Synonyms), and Test 8 (Memory–Paragraph) can be ascribed to the Black subjects' lack of proficiency in the language of the test (English and not their mother tongue). Although Black pupils are taught in English from Standard 3 on, their language skills are not sufficiently developed by the time they have reached Standard 7 to compete with those of Whites.

The suitability of the JAT for use as a common test battery for Blacks and Whites from the point of view of test and item bias is discussed by Owen (1989) and does not directly concern us here. What is of importance, however, is whether Black–White–Indian mean test differences are correlated with the JAT's *g* loadings. To address this problem, we followed Jensen's (1985, pp. 198–200) formulation of the methodology for testing Spearman's hypothesis. The correlation matrices for the 10 tests for each group of pupils (Table 2) were factor analyzed separately by means of the principal-factor method. The first unrotated principal factor is regarded as Spearman's *g*. This factor accounted for 44, 34, and 44 percent of the variance for the White, Black, and Indian samples, respectively, and the second factor for 14, 12, and 14 percent of the variance. Only the first two factors had eigenvalues above unity, and therefore were statistically significant in all three samples. We follow Jensen in interpreting the first factor as Spearman's *g* on the grounds that (a) it explains 34% to 44% of the variance and (b) the test consists of a varied mix of the major primary abilities (i.e., verbal and nonverbal reasoning, verbal comprehension, spatial abilities, mechanical ability, and memory). We believe that in these circumstances, the first factor can reasonably be interpreted as a measure of Spearman's *g*.

For each group, this factor was corrected for attenuation by dividing the loadings of the various tests on the factor by the square root of the reliability coefficient for the test for the group concerned (i.e., White *g* was corrected by means of White KR-20s, etc.). Note that the reliability coefficient was not available for Test 5 (because as a speeded test, the KR-20 is not applicable), therefore, we gave it the mean of reliabilities of the other 9 tests. From Table 3, it is evident that the three corrected principal factors (*g* loadings) are very similar. On the strength of the very high congruence coefficients (0.99) for the White–Indian and White–Black *g* loadings it can be assumed that, to a large extent, the *same* factor is measured in the three groups.

To test Spearman's hypothesis, we calculated correlations between the corrected *g* loadings (Table 3) and the corrected mean test score differences (Table 1). The *g* loadings for the Black sample were correlated with the Black–White test score differences at $r = .624$, $p < .05$, and thus confirmed Spearman's hypothesis. The *g* loadings for the White sample were correlated with the test score

Downloaded by [New York University] at 12:47 06 December 2014

Downloaded by [New York University] at 12:47 06 December 2014

TABLE 2
Intercorrelations of the 10 Test Scores of the JAT for White, Indian, and Black Pupils

Tests	1	2	3	4	5	6	7	8	9	10
					White pupils (n = 1,056)					
JAT 1: Classification										
JAT 2: Reasoning	0.35									
JAT 3: Number ability	0.22	0.63								
JAT 4: Synonyms	0.32	0.61	0.42							
JAT 5: Comparison	0.13	0.36	0.45	0.27						
JAT 6: 2-D	0.39	0.47	0.34	0.34	0.25					
JAT 7: 3-D	0.45	0.49	0.32	0.37	0.23	0.71				
JAT 8: Memory (paragraph)	0.24	0.50	0.42	0.45	0.38	0.23	0.23			
JAT 9: Memory (symbols)		0.19	0.42	0.37	0.34	0.31	0.25	0.27	0.52	
JAT 10: Mechanical insight	0.35	0.52	0.35	0.37	0.21	0.53	0.51	0.31	0.22	
					Indian pupils (n = 1,063)					
JAT 1: Classification										
JAT 2: Reasoning	0.42									

Downloaded by [New York University] at 12:47 06 December 2014

JAT 3: Number ability	0.30	0.64							
JAT 4: Synonyms	0.33	0.67	0.51						
JAT 5: Comparison	0.20	0.40	0.40	0.41					
JAT 6: 2-D	0.49	0.44	0.32	0.27					
JAT 7: 3-D	0.51	0.48	0.34	0.31	0.22	0.73			
JAT 8: Memory (paragraph)	0.16	0.47	0.38	0.49	0.31	0.16	0.18		
JAT 9: Memory (symbols)	0.22	0.44	0.40	0.42	0.31	0.27	0.30	0.43	
JAT 10: Mechanical insight	0.33	0.42	0.34	0.38	0.23	0.45	0.44	0.25	0.29

Black pupils (n = 1,093)

JAT 1: Classification									
JAT 2: Reasoning	0.29								
JAT 3: Number ability	0.26	0.29							
JAT 4: Synonyms	0.23	0.33	0.27						
JAT 5: Comparison	0.28	0.27	0.43	0.25					
JAT 6: 2-D	0.37	0.28	0.26	0.21	0.30				
JAT 7: 3-D	0.38	0.30	0.21	0.18	0.22	0.64			
JAT 8: Memory (paragraph)	0.24	0.28	0.24	0.25	0.32	0.18	0.16		
JAT 9: Memory (symbols)	0.22	0.21	0.28	0.15	0.28	0.25	0.23	0.35	
JAT 10: Mechanical insight	0.26	0.24	0.23	0.22	0.25	0.23	0.21	0.23	0.19

Downloaded by [New York University] at 12:47 06 December 2014

TABLE 3
***G* Loadings of the JAT (Unrotated First Principal Factor) for White, Indian, and Black Pupils Corrected for Attentuation**

JAT	White		Indian		Black	
	g loading	Corrected *g**	*g* loading	Corrected *g***	*g* loading	Corrected *g****
1	.457	0.601	.540	.667	.542	.678
2	.822	0.924	.835	.949	.529	.928
3	.647	0.727	.665	.747	.520	.642
4	.652	0.741	.709	.788	.448	.830
5	.473	0.531	.474	.539	.550	.714
6	.688	0.740	.658	.708	.653	.718
7	.697	0.758	.688	.740	.652	.733
8	.595	0.669	.524	.616	.468	.578
9	.518	0.557	.556	.604	.469	.510
10	.623	0.716	.559	.726	.424	.744

*$g/\sqrt{\text{White KR-20}}$. **$g/\sqrt{\text{Indian KR-20}}$. ***$g/\sqrt{\text{Black KR-20}}$.

differences at $r = .235$, but this correlation was not statistically significant. It may seem curious that there should be such a large difference between the two correlations despite the very high coefficient of congruence between the *g* loadings. To examine this discrepancy further, we calculated Spearman's rank order correlations between the *g* loadings (White) and the Black–White test differences ($r = .20$, *ns*), the *g* loadings (Black) and the Black–White test differences ($r = .64$, $p < .05$), and the Black *g* loadings by the White *g* loadings, $r = .72$, $p = .05$).

These correlations show that the rank ordering of the *g* loadings and White–Black test differences were not the same for the two groups. Bearing in mind that we used only 10 tests, it stands to reason that one or two aberrant *g* loadings and/or test differences could sway the results in one direction or another. An inspection of the data in Table 3 revealed that Test 5 had the lowest *g* loading of all for the White sample (this was to be expected because the test primarily requires perceptual speed) but an unexpectedly high *g* loading for the Black sample. The mean test score differences between Whites and Blacks for this test (Table 1), on the other hand, ranked only fifth. If the rank orders of the two variables (*g* loadings and White–Black test differences) had been the same in just two instances, namely, Test 5 and Test 1, the Spearman *r* for Whites would have been .65 (the same as that for Blacks), instead of .23. Under these circumstances the Spearman hypothesis would also have been supported by data for the Whites.

We thought it would also be interesting to examine the Indian–White test score differences to see whether they were a function of the *g* loadings. As with the Black–White comparison, this analysis can be done for the *g* loadings for the Indian and for the White samples. For the Indian sample, the correlation was +.129; for the White sample, it was +.081. Both correlations were negligible, showing that Indian–White test score differences have no relation to differences in Spearman's *g*. Of course, neither Spearman nor Jensen has suggested that there would be any relationship. What makes this result interesting is that it shows that population differences in intelligence are not necessarily primarily differences in Spearman's *g*, as Sternberg (1985) suggested, or even statistical artifacts, as argued by Schonemann (1985). Evidently, it is possible to find population differences of 1 *SD* unit that do not, to any significant extent, consist of differences in Spearman's *g*.

REFERENCES

Irvine, S. H. (1969). Figural tests of reasoning in Africa. *International Journal of Psychology, 4,* 217–228.

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

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.

Downloaded by [New York University] at 12:47 06 December 2014

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

Osborne, R. T., & McGurk, F. C. J. (Eds.). (1982). *The testing of Negro intelligence*, Vol. 2. Athens, GA: Foundation for Human Understanding.

Owen, K. (1989). *Test and item bias: The suitability of the Junior Aptitude Tests as a common test battery for White, Indian and Black pupils in Standard 7.* Pretoria: Human Sciences Research Council. (ERIC No. TM 013999).

Schonemann, P. H. (1985). On artificial intelligence. *Behavioral and Brain Sciences, 8,* 241–242.

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

Sternberg, R. J. (1985). The Black–White differences and Spearman's *g:* Old wine in new bottles that still doesn't taste good. *Behavioral and Brain Sciences, 8,* 244.

Verwey, F. A., & Wolmarans, J. S. (1983). *Junior Aptitude Tests.* Pretoria: Human Sciences Research Council.

Received March 1, 1993

Downloaded by [New York University] at 12:47 06 December 2014