Intelligence 39 (2011) 456–467

Contents lists available at SciVerse ScienceDirect

Intelligence

The Flynn effect in South Africa

Jan te Nijenhuis [a,*,1], Raegan Murphy [b,1], Rene van Eeden [c]

[a] Work and Organizational Psychology, University of Amsterdam, The Netherlands
[b] Applied Psychology, University College Cork, Ireland
[c] Department of Psychology, University of South Africa, South Africa

ARTICLE INFO

Article history:
Received 4 March 2011
Received in revised form 12 August 2011
Accepted 13 August 2011
Available online 1 September 2011

Keywords:
Flynn effect
Secular score gains
g
IQ tests
Intelligence
South Africa

ABSTRACT

This is a study of secular score gains in South Africa. The findings are based on representative samples from datasets utilized in norm studies of popular mainstream intelligence batteries such as the WAIS as well as widely used test batteries which were locally developed and normed in South Africa. Flynn effects were computed in three ways. First, studies where two different groups take the same test, with several years in between, using representative or comparable samples were used. Second, studies where the same group takes two different test batteries at a specific time were used. Third, the score differences between English- and Afrikaans-speaking Whites in South Africa in the 20th century were compared. The Flynn effect in White groups in South Africa is somewhat smaller than the Flynn effect in Western, industrialized countries (total $N = 6534$), and the Flynn effect in Indian groups is substantially smaller (total $N = 682$). Non-verbal IQ scores surpassed increases in verbal IQ scores. The findings from English- and Afrikaans-speaking Whites evidence a leveling out of differences in score gains over the 20th century (total $N = 79{,}310$). A meta-regression analysis showed no clear support for the moderators a) method used for computing the Flynn effect gain, b) type of test battery, c) time span, d) quality of the sample, and e) average age of sample.

1. Introduction

Secular score gains in IQ test scores are one of the most intriguing and controversial findings in the recent psychology research literature. James Flynn (1984, 1987) was the first to show that average scores on intelligence tests are rising substantially and consistently, all over the world. These gains have been going on for the better part of a century — essentially ever since standardized tests were developed. Although Flynn effects have been shown for many countries, as yet, little has been done on the Flynn effect in South Africa. This paper is the first to describe the Flynn effect in South Africa using a large number of primary studies.

The Flynn effect refers to the increase in IQ scores over time. For Western, industrialized countries, between 1930 and 1990 the gain on standard broad-spectrum IQ tests averaged three IQ points per decade. This trend has continued to the present day in the United States (Flynn, 2007, 2009a). In the United Kingdom, gains on the Raven's Progressive Matrices are still robust except oddly, at ages 13 to 15, an exception confirmed by Piagetian data (Flynn, 2009b; Shayer, Ginsburg, & Coe, 2007). It is a global phenomenon and has been recorded for a number of industrialized and non-industrialized nations including countries in Africa (Flynn, 2006). For verbal tests, or more precisely, tests with a content that most reflects the traditional classroom subject matter, the gain is 2 IQ points per decade, and for non-verbal (fluid and visual) tests the gain is 4 IQ points per decade (Jensen, 1998).

Recently, however, studies from Denmark, Norway, and Britain show that the secular gains have stopped and even suggest a decline of IQ scores (Lynn, 2009; Shayer et al., 2007; Sundet, Barlaug, & Torjussen, 2004; Teasdale & Owen, 2008).

* Corresponding author at: University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. Tel.: +31 20 6002067.
E-mail address: JanteNijenhuis@planet.nl (J. te Nijenhuis).

[1] The first two authors contributed equally; the order of names is random.

0160-2896/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2011.08.003

However, there is also recent evidence of IQ test scores continuing to rise in Western, industrialized countries (e.g. in France, see Bradmetz & Mathy, 2006) and in countries in the former communist Eastern Europe (e.g. in Estonia, see Must, te Nijenhuis, Must, & van Vianen, 2009). Recent studies show IQ scores rising in less-developed parts of the world, for example in Kenya (Daley, Whaley, Sigman, Espinosa, & Neumann, 2003), Sudan (Khaleefa, Abdelwahid, Abdulradi, & Lynn, 2008) and in the Caribbean (Meisenberg, Lawless, Lambert, & Newton, 2006). However, there are, to this date, only two studies of the Flynn effect in South Africa (Richter, Griesel, & Wortley, 1989; Wicherts, Dolan, Carlson, & van der Maas, 2010).

Various causes have been hypothesized for the Flynn effect, including education, nutrition, health care, inbreeding, GDP, urbanization, family size, health care expenditure, the dissemination of visual–spatial toys, and teacher to student ratio (see Flynn, 2006, 2007; see Jensen, 1998, ch. 10). It is difficult to conclude what the most important cause is, as many of the effects take place at the same time and show similar trends.

1.1. Racial classification and segregation in education in South Africa

The four racial groups currently classified in the country are Black, White, Colored, and Indian. The latest South African Government statistics (2007 mid-year) reveal a total populace of 47,850,700 of which Blacks account for 79.6%; Whites account for 9.1%; Coloreds account for 8.9% and Indian/Asian account for 2.5%. Whites are of European descent and a distinction is made between Afrikaans- and English-speaking Whites. About 60% of the White population of South-Africa are Afrikaans-speakers. The Afrikaans-speaking are chiefly descendant from the French Huguenots and Dutch peoples. Historically their social development sprang from an impoverished rural base (Claassen, 1997). In 1946 the per capita income of Afrikaans-speaking Whites was 47% of that of the English-speaking Whites; in 1960 it was 60%, and in 1976 it was 71%. About 40% of the Whites are English-speaking and traditionally they completed more years of secondary and tertiary schooling, but this has changed through the years from the early twentieth century and both English- and Afrikaans-speaking Whites are more or less on a par (Claassen, 1997). Throughout the 19th century the majority of Afrikaans-speakers were educated at home, due in part to the resistance of English education for their children in the British-dominated education system and enrollments among this group was considerably lower than enrollments of English-speakers in the British-held territories (Byrnes, 1996). The effects of these choices were felt throughout half of the 20th century in the South African education system. The increase in number of years of education through the years has been hypothesized to partly account for the increases evidenced in IQ scores for the latter group (Claassen, 1997; see also Ceci, 1991 and Jensen, 1998). Coloreds are of mixed racial origin spanning numerous countries outside Africa but having substantial genetic Southern African ancestry (some Coloreds are of Bantu–Khoisan descent). This term does not have the same meaning as the American term 'Colored'. In South Africa it does not refer to a Black person. The reason for the presence of Indian populations is that in the nineteenth century the European colonists needed laborers for manual work of various kinds. Indians were brought over from the 1860s onwards principally to work in the sugar and cotton plantations in Natal. It must be recalled that during the Apartheid era, national education was decentralized regarding access to equal opportunities and resources. Education for Black school children was by and large severely below the White counterpart standards (Shuttleworth-Edwards et al., 2004). Difference in schooling is also reflected in the difference between White and non-White access to higher education. In 1921, 22% of the White population attended public or private schools in comparison to 4.6% of non-Whites. However, in 2007, 76% of the non-White population attended public or private schools and 73% of the White population attended private or public schools in 2007 (Department of Education, 2008). As of 2006, non-Whites make up 74% of the student population attending public higher education institutions and similarly make up 77% of the student population conducting distance education (Department of Education, 2008).

1.2. IQ testing and group differences in South Africa

During the early part of the 20th century South African test developers utilized existing international test batteries as their main source of test information usually derived from the Binet type individual test and the Army Beta group test (Fick, 1929, 1939). As early as 1916, the Moll–Leipoldt Scales had been compiled, initially under the name 'Binet–Simon–Goddard–Healy–Knox Scale' with a group intelligence test being released in 1924 at the University of Stellenbosch (Smit, 1996). Through the intervening years (1924–2008) a number of international tests were utilized and/or standardized for local South African conditions. South African-developed tests include, among others, the South African Group Intelligence Test (SAGIT), and the Individual Scale of General Intelligence for SA (ISGIS). The testing tradition in South Africa thus reflects an amalgamation of original uniquely developed and normed tests as well as normed and locally standardized international tests (Huysamen, 1996).

South African literature has shown for many decades that substantial differences in test scores between various cultural and language groups exist (Biesheuvel & Liddicoat, 1959; Claassen, Krynauw, Paterson, & wa ga Mathe, 2001; Dent, 1949; Fick, 1929; Foxcroft & Aston, 2006; Irvine, 1969; Knoetze, Bass, & Steele, 2005; Rushton, 2001; Rushton & Skuy, 2000; van der Berg, 1989; Verster & Prinsloo, 1988). The most cited, influential, and detailed book on group differences in IQ is by Lynn and Vanhanen (2002). Based on a review of the South African literature on IQ testing (pp. 218–219) they estimate the IQs for four groups: Whites: 94; Coloreds: 82; Blacks: 66; and Indians: 83. Whites outscore non-Whites, and within the White group English-speakers outscore Afrikaans-speakers; the latter finding has been well documented (Mariotti, 2009). Increasing access to education for the Afrikaans-speakers was a direct result of the social policies implemented from the mid-1940s onward; this is most likely the most important factor (Prinsloo, 2007). The Nationalist Party, which came to power in 1948, instituted an affirmative action program to enable Afrikaans speakers greater participation in the labor force. This continued till the mid 1960s at which time Afrikaans-speakers were more settled in urbanized life (Louw, 2004). Often, positions were reserved for poor Afrikaans-speakers in the state-owned railways and civil service. These were generally low-paid jobs but provided stable home environments 'for children entering the expanding Afrikaans education system' (Louw, 2004, p.57). As part of these policies Afrikaans-language schools and universities came into existence. Socio-cultural reasons are also cited as contributing to historical differences in cognitive performance between English- and Afrikaans-speakers. For instance, the latter group are known, in the past, to have been more nationalistic, conservative, authoritarian, religious and group-oriented when compared to the English-speakers (Katz, 1988; Prinsloo, 2007). The relation between authoritarianism and cognitive performance has been identified in these two cultural groups and it has been suggested to be causal (Duckitt, 1992). However, it has also been long known and cited that socio-economic status, educational attainment, language bias, socio-political circumstances, and test familiarity play a role in the depressed scores of certain groups even though any one of these factors cannot be solely accountable for group differences (Biesheuvel, 1952a,b; Biesheuvel & Liddicoat, 1959; Crawford-Nutt, 1976, 1977; Furnham, Mkhize, & Mndaweni, 2004; Kamin, 2006; Liddicoat & Roberts, 1962; Lynn & Owen, 1994; Owen, 1992; Pressey & Teter, 1919; Rushton, 2008; Shuttleworth-Edwards et al., 2004; Skuy et al., 2002; van der Berg, 1989; van Eeden, de Beer, & Coetzee, 2001).

Due to its ethnic diversity, the large differences between groups on many variables, the availability of high-quality psychometric tests, and an extensive literature on testing, South Africa seems an almost ideal country to study secular score gains. However, the unique manner of sampling in post-democracy South Africa resulted in different groups being clustered together such that Whites, Coloreds, Indians, and Blacks are taken as one group reflecting an overall 'South African' IQ score. This manner of sampling was strongly dependent on the SES group to which individuals were assigned. In contrast, pre-democracy South African sampling often stratified the samples according to different race groups such that separate IQ scores were available for the different groups. Added to this mélange of sampling mixes were group clusters of Whites and Coloreds, Indians and Blacks, or Whites and Indians. This makes it difficult to impossible to disentangle the different groups' separate scores, which has as a consequence that notwithstanding the wealth of South African data, only a small percentage could be used in the present analyses of secular score gains. More specifically, due to the absence of good datasets on Black South Africans our study was limited to Indian and White South Africans.

1.3. Research questions

The first research question is whether the secular gains for total IQ, performance IQ, and verbal IQ of White South Africans is comparable to that found in Western, industrialized countries. The second research question focuses on the size of the secular gain for Indian South Africans. The third question is how the differences between English-speaking and Afrikaans-speaking White South Africans compare in the 20th century.

2. Method

2.1. Data gathering

The data were gathered by two of the authors, who both have extensive knowledge of cognitive testing in South Africa and published extensively on the topic. There is no one single test repository in South Africa. Moreover, the Human Science Research Council's test library has been disbanded making it difficult to gain access to the collection of manuals. Searchable South African data archives such as Sabinet yielded part of our database of results. Manuals from test libraries housed at some universities in South Africa were also searched by two of the authors. A number of published articles, from which data was sourced, are only available in hard-copy format and in some instances can only be found in South Africa. Some results were sourced from postgraduate dissertations and are only available in Afrikaans. Other data-sets were sourced from internationally indexed research. Although we did not aim to find every last relevant study, within the restrictions described above, we believe our search was near exhaustive.

The test batteries that were sourced for this paper were standardizations of large-scale testing in the country. For instance, the New South African Group Test was standardized on representative samples ranging in size from 4448 to 6486. Due to the need for psychometric equivalence across cultural groups, sampling was often quota-based for each of the four groups. The batteries sourced were normed and standardized psychometric tests either developed in South Africa or imported and locally standardized.

2.2. Tests

A brief review of all tests used in the computation of a Flynn effect is given in Table 1. The tests are listed chronologically and according to date of data gathering. Older material was less researched and information in terms of sample descriptions is limited in certain instances. The subtests included in the individual and group tests listed here resemble those included in international batteries commonly used.

2.3. Samples

In total 12 test batteries were sourced. The database was subdivided into 87 sets of data which were subsequently analyzed for use in the study but not all were ultimately utilized due to missing data or data that could simply not be compared over the years due to the nature of the sampling (early sampling according to race as opposed to later sampling according to SES, as described earlier). The data-sets used for this paper contain IQ scores gathered between 1925 and 2000 and are composed of samples of individuals born between 1890 and 1985. Sample sizes range from 24 to 40,000 depending on the nature of the assessment (small-scale research or standardization).

2.4. Statistical analyses

2.4.1. Computation of the secular score gains

In this paper three methods of computing secular gain scores were used.

1) The first method was used in Flynn (1987). A comparison was made between the outcomes from studies using the same test in different groups, with at least five years in between, in all cases using representative or comparable samples (i.e. in terms of age, sex, population group, education, etc.). For instance, the Raven's Progressive

Matrices was taken in both 1960 and 2000 by samples of comparable groups. The score increase is an estimate of the Flynn effect.

2) The second method was used in Flynn (1984). In studies where the same group took two different test batteries the resulting means were compared. There had to be at least four years between the norm samples of the two tests. These samples need not be representative. For instance, one group took both the SSAIS-R (1987) and the NSAGT (1954). For instance, if the same group of subjects took the NSAGT – normed in 1954 – and the SSAIS-R – normed in 1987 – they should score higher on the earlier test, the NSAGT. The group's raw score on the NSAGT should be compared to the norm scores of the NSAGT from 1954, which might result in a score of 107. The group's raw score on the SSAIS-R should be compared to the norm scores of the SSAIS-R from 1987, which might result in a score of 101. The difference between their mean scores on the two tests serves as a measure of the magnitude of gains, that is, scoring 107 on the earlier test and 101 on the later test suggests a gain of 6 IQ points in 33 years.
3) Using data going back to people born in the 1890s Verster and Prinsloo (1988) and Claassen (1997) describe how the English-speaking outscore the Afrikaans-speaking and how the erstwhile large gap diminished within four of five generations. Many of the studies they cite used carefully collected, representative samples. Although the aforementioned authors' results were not used by them to test the Flynn effect, these data can be used to estimate the size of the Flynn effect for the Afrikaans-speaking group. We use a three-step procedure: first, from people born in 1890 to people being born in 1985 the English-speaking means are compared with the Afrikaans-speaking means; second, using the results from the two estimation methods described above gives a clear estimate of the score gains for the English-speakers; third, combining the gains from the first step and the second step results in an estimate of the score gains for the Afrikaans-speaking.

So, in a sense the scores of the English-speaking are used as a yardstick, albeit that the yardstick is not disconnected from the Flynn effect. Another way to look at it, is to think of how the Afrikaans-speaking catch up, by comparing their scores with the English-speaking from people being born in 1890 to people being born in 1985.

In the test–retest method described here it is essential that the test order is counter-balanced. Detailed information on this topic cannot always be gleaned from the original studies, but South Africa for decades had one of the best testing traditions in the world. All studies were done by professional researchers, including one of the authors of the present paper, who would have taken this well-known effect into account.

2.4.2. Testing moderator variables

The gain scores could be influenced by the methodology used, the test battery used, and the time span covered. So, we tested the influence of these and other moderator variables. We took all the data points from the White samples from the present study. We decided not to combine them with the two data points from two previous studies, because they used, respectively, the Raven's and the Draw-a-Man, whereas the present study included only data points based on test batteries. This resulted in all data on the Flynn effect on White samples using test batteries in South Africa. These data were then analyzed using meta-analytical techniques.

There are several ways to carry out meta-analyses (Hunt, 1997), and we choose the Hunter and Schmidt style (Hunter & Schmidt, 1990, 2004) meta-analysis. In the first step all data are used in a bare-bones meta-analysis – correcting only for sampling error – and then in a second step subsets are created, based on potential moderators; on these subtests separate meta-analyses are carried out. Hunter and Schmidt (1990, pp 292–303) argue that when testing for moderators it is preferable to carry out separate meta-analyses than to correlate the effect size with the moderator. The Schmidt and Le (2004) program is based on a random effects model, so compared to traditional weighted means large samples get smaller weights and small samples get larger weights.

So, in Step 1 we collected all effect sizes (d) and we corrected for differences in sample size. Step 1 results in a mean value of the effect size and a percentage of variance explained in the data points by sampling error. When the percentage of variance explained was similar to or higher than 75 we did not look for moderators (see Hunter & Schmidt, 1990). The '75% rule' states that if 75% or more of the observed variance across studies is due to statistical artifacts, then probably all of it is artifactual variance, on the grounds that the remaining 25% is likely due to artifacts not corrected for. Thus, in cases where 75% or more of the variance is explained by artifacts, including sampling error variance, moderators are unlikely to have caused a real variation in observed correlations.

When the percentage of variance explained in the data points was smaller than 75 we continued to Step 2, where we tested the influence of moderators. This meant breaking up the sample in at least two subsets and performing separate meta-analyses on them. Hunter and Schmidt (1990, p. 293) state that there is a moderator when 1) the difference in the size of the effect between groups is theoretically meaningful, and 2) there is a reduction in unexplained variance. Hunter and Schmidt strongly advice to use moderators that make theoretical sense. The moderators tested were:

a) method used for computing the Flynn effect gain: the same test used in different group versus two tests given to the same group
b) type of test battery; the Flynn effect is generally largest for fluid and visual subtests of a battery. So, if the first test battery has little of these subtests and the second test battery has many of these subtests one would expect the largest increase. We computed for the first test battery the percentage of fluid and visual subtests; we computed for the second battery the percentage of fluid and visual subtests; then we computed $d =$ (% second test battery) − (% first test battery).
c) time span
d) quality of the sample: were data from the second testing compared to a standardization sample or were they compared to another kind of sample
e) average age of sample.

Time of data sampling could not be used as a moderator, because all the samples were collected in the 1980s.

Table 1
South-African and international test batteries used in the present study, full names, dates of issue of test manual, and date the standardization sample was collected.

Test battery	Full name	Date[a]		Age groups	Norm groups
		Manual	Stand. sample		
OSAIS	The individual scale of the National Bureau for Educational Research (Old South African Individual Scale); released in 1937 and partly based on the Stanford–Binet Scale of 1916	1939	1937		
NSAIS also named SSAIS	The New South African Individual Scales or Senior South African Individual Scale	1964/1970 name change 1980	1962	5–17	1590 Afrikaans-speaking White and 812 English-speaking White children
SSAIS-R	Senior South African Individual Scale—Revised	1991	1987	7 y–16 y 11 mo	2000 White, 2000 Colored and 2000 Indian
SAGIT	South African Group Intelligence Test	1933[b]	1931	10–16	Forms A1 and A2 for Afrikaans-speakers; forms E1 and E2 for English-speakers
OMHIS	Official Mental Hygiene Individual Scale	1929	1927		1500 randomly selected from a population of 10,000 pupils
ISGIS	Individual Scale of General Intelligence for SA	1939	1937[b]		
JSAIS	Junior South African Individual Scale	1979	1976	3–7	1795 stratified sample
GTISA junior	Group test for Indian South Africans	1968	1966		Standardized on Indian pupils
GTISA intermediate	Group test for Indian South Africans	1983	1981		Standardized on Indian pupils
GSAT junior	General Scholastic Aptitude Test	1990	1989	9 y 0 m–11 y 11 m	Representative of the White, Colored, and Indian populations
GSAT intermediate	General Scholastic Aptitude Test	1987	1984	11 y 0 m–14 y 11 m	Representative of the White, Colored, and Indian populations
GSAT senior	General Scholastic Aptitude Test	1991	1989	14 y 0 m–18 y 6 m	Representative of the White, Colored, and Indian populations
NSAGT junior	New South African Group Test	1965	1951–1956; 1965		Standardized on White school children
NSAGT intermediate	New South African Group Test	1963	1951–1956; 1963		Standardized on White school children

NB group test junior	National Bureau Group Test for White pupils	1974	1972[b]	11–13	
NB group test intermediate	National Bureau Group Test	1974	1970	13–15	A stratified norm group of 3123 white pupils
NB group test senior	National Bureau Group Test	1974	1971	15–17	2581 white pupils
NB group test 5/6 and 7/8	National Bureau Group Test for 5 and 6 year olds	1960 and the 7/8 year olds 1982 and renormed 1993	1960	5–8	A stratified random sample of 3705
JAT	Junior Aptitude Test	1961/1975	1972[b]	12–16	Standardized on White school children
CPI	Cape Province Individual Scale for Afrikaans-speakers	1929	1925–1927	8–17	Afrikaans-speakers in the Cape
ISGSA	The Individual Scale for General Scholastic Aptitude	1998	1991–1992	4–16	3099 White and Colored pupils. Weighting was used to ensure proportional representation of education departments
International test batteries standardized in South Africa					
SAWAIS	South African Wechsler–Bellevue Adult Intelligence Scale	1962	1958	18–59	2761 volunteers
WAIS III	Wechsler Adult Intelligence Scale III	2001	1998	16–69	1300; all four race groups (25% from each group)
Griffiths	Griffiths Mental Development Scales (translated)	1986	1970[b]		

[a] Dates refer to date of testing (manual dates differ widely in terms of reprints).

[b] Year estimated. When the date at which standardization was carried out is not given, it was assumed to have taken place two years before the date of publication. When the collection of the standardization sample took two years we rounded off to the earliest year, when the collection took three years we took the year in the middle, and when it took four years we took as the date the second year. In older texts the SSAIS is also referred to as the NSAIS (New South African Individual Scale).

Table 2
Comparable White groups taking the same test with several years in between by test battery.

Test	Year born	Year sample	N	Gap	IQ score	Gain pd
NSAGT junior 1965 (van Eeden & Visser, 1992)	1971–1980	1987	556	22 years	106.59	2.99
NSAGT intermediate 1963 (van Eeden & Visser, 1992)	1971–1980	1987	494	24 years	109.03	3.76
NSAGT senior 1965 (van Eeden & Visser, 1992)	1971–1980	1987	136	22 years	105.25	2.38
NSAGT senior 1965 (Claassen, 1983)	1968	1981	2053	16 years	98.75	−0.78

Two of the authors of the present paper judged the values of the moderators for each data point. The most direct gage of the reliability of judges' ratings is the mean correlation between judges' rating of a given moderator. These judgments can only be used when the correlation is high. So, if the first round of ratings of the moderators did not result in a value of one, we discussed until we reached consensus.

3. Results

Table 2 shows the studies where two different but comparable groups took the same test, with several years in between, using representative or comparable samples. The gain per decade for Whites is on average 2.08 IQ points.

Table 3 lists the studies where the same group took two different test batteries at a specific time. The gain per decade for Whites is on average 3.63 IQ points. The gain per decade for Indians is on average 1.57 IQ points.

When the results from Tables 2 and 3 are combined the following figures emerge: The gain per decade for Whites is on average 2.85 IQ points, and the gain per decade for Indians is 1.57 points. On average, the gain score for Whites is somewhat lower than the three points that have been reported in the literature for other industrialized countries (Flynn, 2007). It should be noted that most of the broad test batteries used in the South African samples are similar in content to those in the many other studies on the Flynn effect. The gain for Indians was substantially smaller than the gain for Whites.

Differences between English- and Afrikaans-speaking samples throughout the decades were also compared. Early studies delineated the language groups strictly according to home language spoken among Whites only whereas the later studies included all cultural groups whose first language was English or Afrikaans. So, many of the later studies could not be used for our analyses. The earliest data detailing English-speaking and Afrikaans-speaking differences emanates from the 1950s with a study utilizing the South African Group Test with normed data gathered in 1931 (Smit, 1996; Verster & Prinsloo, 1988). Data from this point forward consistently evidenced a substantial discrepancy between the language groups with higher IQ's being established for the English-speaking groups.

Table 4 shows the score differences between the two groups. The Table is ordered according to date of sample collection. Fig. 1 reports the same data points and clearly shows how the groups are slowly converging in their mean scores. Score differences are computed as: mean of the English-speaking group minus the mean of the Afrikaans-speaking group. A positive score difference means therefore that the English-speaking group has a higher mean score, and a negative score means that the Afrikaans-speaking group has a higher mean score.

Fig. 1 clearly shows that when using only robust datasets the huge score gap between Afrikaans speakers and English speakers strongly diminishes over the run of a century. Fig. 2 shows that when using all the samples the overall picture is very much the same, with the exception of a few outliers. So,

Table 3
Same groups taking different tests at a specific time.

Test	Global IQ scores	N	Gap	Gain per decade
Sample of Whites tested in 1987 aged 7–16 born 1971–1980				
NSAGT junior (1965) and SSAIS-R (1987) (van Eeden & Visser, 1992)	106.59 and 101.59	556	22 years	2.27
NSAGT intermediate (1963) and SSAIS-R (1987) (van Eeden & Visser, 1992)	109.03 and 102.48	494	24 years	2.72
Sample of Whites tested in 1987 aged 6–14 born 1973–1981				
OSAIS (1937) and JSAIS (1976) (Robinson & Boshoff, 1990)	104.8 and 95.8	237	39 years	2.3
OSAIS (1937) and SSAIS (1962) (Robinson & Boshoff, 1990)	104.8 and 108.5	421	25 years	−1.48
SSAIS (1962) and JSAIS (1976) (Robinson & Boshoff, 1990)	108.5 and 95.8	237	14 years	9
OSAIS (1937) and NB 5/6 (1960) (Robinson & Boshoff, 1990)	104.8 and 107.7	62	23 years	−1.26
OSAIS (1937) and NB 7/8 (1960) (Robinson & Boshoff, 1990)	104.8 and 109.5	225	23 years	−2.04
OSAIS (1937) and GSAT intermediate (1984) (Robinson & Boshoff, 1990)	104.8 and 93.2	160	47 years	2.46
NB 5/6 (1960) and JSAIS (1976) (Robinson & Boshoff, 1990)	107.7 and 95.8	237	16 years	7.43
NB 7/8 (1960) and JSAIS (1976) (Robinson & Boshoff, 1990)	109.5 and 95.8	237	16 years	8.56
SSAIS (1962) and GSAT intermediate (1984) (Robinson & Boshoff, 1990)	108.5 and 93.2	160	22 years	6.95
OSAIS (1937) and JSAIS (1976) (Robinson & Boshoff, 1990)	104.8 and 95.8	237	39 years	2.3
Sample of Whites tested in 1982 aged 3–7 born 1975–1979				
Griffiths (1970) and JSAIS (1976) (Luiz & Heimes, 1994)	105.81 and 100.97	32	6 years	8.07
Sample of Indians tested in 1989 aged 15 born 1973:				
GTISA junior (1966) and GTISA intermediate (1981) (Claassen et al., 1991).	110.71 and 108.35	158	15 years	1.57

Table 4
English and Afrikaans differences ordered by year of publication of study and according to year of birth with positive group differences denoting higher scores for English-speaking.

Study	Average year born	Year sample	Test	*N* Afr.	*N* Eng	Diff verbal	Diff perf	Diff IQ
Olckers (1950)[a]	1938	1950	S. A. Group Test	630	1170	n.r.	n.r.	7
Morkel (1950)[a]	Approx 1932	1950	Mental alertness	500	502	n.r.	n.r.	1.8
Biesheuvel and Liddicoat (1959)[b]	1896	1950	SAWAIS	45	68	14.78	17.04	16
	1901	1950	SAWAIS	110	86	12.71	11.87	12.3
	1906	1950	SAWAIS	138	99	10.32	12.95	11.6
	1911	1950	SAWAIS	175	120	5.95	5.85	5.9
	1916	1950	SAWAIS	226	149	7.39	8.18	7.8
	1921	1950	SAWAIS	222	148	9	8.74	8.9
	1926	1950	SAWAIS	227	152	5.32	8	6.7
	1931	1950	SAWAIS	240	160	6.62	8.76	7.7
	1936	1950	SAWAIS	240	156	7.92	9.21	8.5
Biesheuvel (1952b)	1930? (not stated)	1950	RPM	n.r.	n.r.	7.5		
Langenhoven (1957)[a]	1941? (not stated)	1954	NSAGT	n.r.	99	n.r		
TALENT[b,a]	1952	1965	NSAGT	40,900	21,129	5.17	7.44	6.34
TALENT[b,a]	1952	1965	JAT	40,767	21,083	n.r	n.r.	2.77
TALENT[b,a]	1952	1967	SAT	7071	4719	3.18		3.18
Cudeck and Claassen (1983)	1969	1981	NSAGT-G	171	319	n.r.	n.r.	5.0
Claassen (1983)	1968	1981	NSAGT Int.	786	1266	5.77	−2.4	1.5
Luiz and Heimes (1994); Robinson (1994)	1974	1981	JSAIS	90	32	0.15	2.67	1.24
Claassen (1990)[b]	1970–1972	1984	GSAT	215	299	4.21	4.05	3.12
GSAT manual (1990)[b]	1977	1988	GSAT junior	1963	1635	1.73	3.06	2.52
van Eeden (1991)[b]	1970	1987	SSAIS-R	2967	1709	5.7	5.1	5.25
Claassen et al. (2001)	1929–1984	1999	WAIS-III	97	70	4.97	1.1	3.04[c]

Morkel (1950) does not give the effect size, but reports that the effects are significant, so we conservatively choose a value of 0.05 for the significance coefficient. We computed the effect size, using the formula
$d = \sqrt{f}(n1+n2)/n1xn2)(n1+n2)/n1+n2-2)$. $F(1, 1002)$ at $p<0.05$ yields a value of 3.84 (f is based on the two degrees of freedom: sample size and number of groups). Therefore $\sqrt{3.84(0.003992)(1.002)} = 0.12$ *SD*.
W-B = Wechsler Bellevue; NSAGT Int. = NSAGT intermediate. Biesheuvel and Liddicoat (1959) report data separated for males and females, which we combined. On page 49 of Claassen's (1983) document he states that for his sample E, he cannot be sure how representative the sample is, because it is only representative of school-going 13 year olds in urban areas who are White. Therefore we did not use this subsample for our computations.
[a] Cited in Verster and Prinsloo (1988).
[b] Denotes a representative data set.
[c] For the study of Claassen et al. (2001) the difference in IQ was computed as the mean of the difference in verbal and the difference in performance.

the quality of the datasets does not seem to strongly influence the conclusions. This means that the secular score gain is stronger for the Afrikaans speakers than for the English speakers.

We also investigated that the IQ score increases at the subtest level for the group for which we had data available, in this case for Whites. Table 5 uses the datasets of Table 2 and Table 6 uses data of Table 3, but unfortunately, not all studies report information at the subtest level. It can be clearly seen that non-verbal IQ scores have increased more so than verbal IQ scores and this is in keeping with the literature (Flynn, 2007; Jensen, 1998).

When the mean gains per decade from Tables 5 and 6 are combined the following results emerge: the verbal IQ gain per decade for Whites is on average 2.43 IQ points and the non-verbal gain per decade is 4.75 IQ points (the gain is 0.98 for verbal and 3.78 for non-verbal from Table 5 and 3.88 for verbal and 5.72 for non-verbal from Table 6).

3.1. Testing moderator variables

Psychometric meta-analytical techniques (Hunter & Schmidt, 1990, 2004) were applied using the software package developed by Schmidt and Le (2004). In many cases sampling error explains the majority of the variation between studies, so the first step in a psychometric meta-analysis is to correct the collection of effect sizes for differences in sample size between the studies. The data points and their average sample sizes are reported in Tables 2 and 3. It is clear that there is a large variety of sample sizes. The data derived from seventeen data points with participants totals 6534.

Table 7 presents the results of the psychometric meta-analysis of the seventeen data points (from Tables 2 and 3). It shows (from left to right): the number of correlation coefficients (K), total sample size (N), the mean observed Gain Per Decade (GPD) and their standard deviation (SD_{GPD}). The next two columns present the percentage of variance explained by artifactual errors (%VE), and the 80% confidence interval (80% CI). This interval denotes the values one can expect for the true effect size in sixteen out of twenty cases. The analysis of all seventeen data points yields an estimated Gain Per Decade of 1.90 IQ points, with only .2% of the variance in the observed correlations explained by sample size. This is an extremely low percentage of variance explained suggesting the existence of 1) one or more powerful moderators, 2) outliers and extreme outliers, or 3) both. So, we continued to Step 2, where we tested for the moderators a) method used for computing the Flynn effect gain, b) type of test battery, c) time span, d) quality of the sample, and e) average age of sample. As Table 7 clearly shows there is no clear support for moderators: although the size of the GPD varies in a theoretically meaningful way, the percentage variance explained does not increase in all but one comparison. Even time span cannot be considered a clear moderator: although all the variance is explained when there is

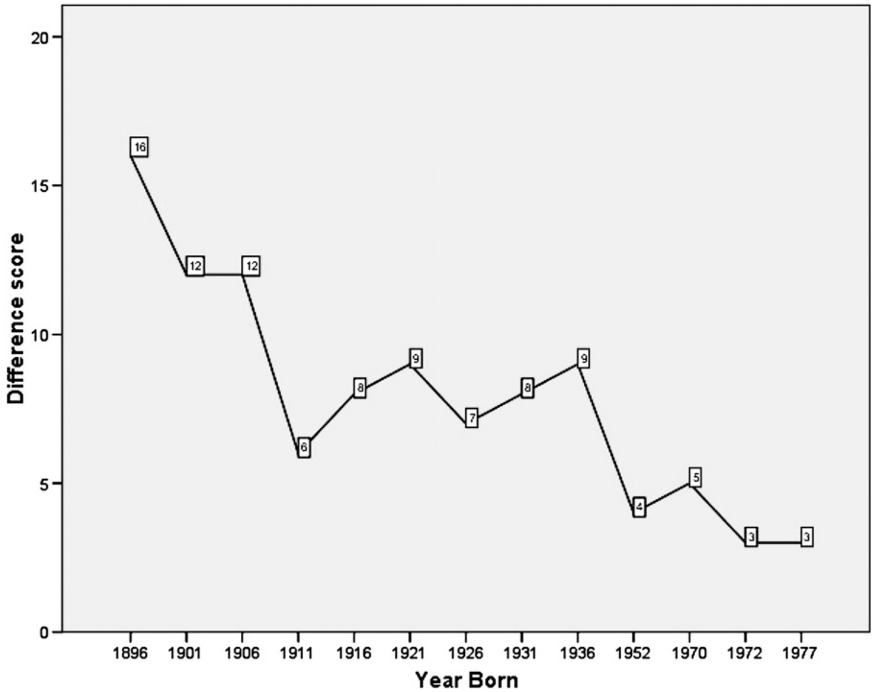

Fig. 1. Score differences between English-speaking and Afrikaans-speaking South-Africans using only robust data-sets samples. Note. We used all representative datasets: the data from Biesheuvel and Liddicoat (1959); the Talent data for 1952 and 1965 are not independent, so we choose the data using the NSAGT – a classical IQ test – over the data using the JAT — which is an aptitude test; the NSAGT also has the largest sample size; the Talent data for 1952 and 1967; the data from the GSAT manual Claassen, Van Niekerk, and Kotzé (1990); the data from Claassen, De Beer, Hugo, and Meyer (1991); and the data from van Eeden (1991).

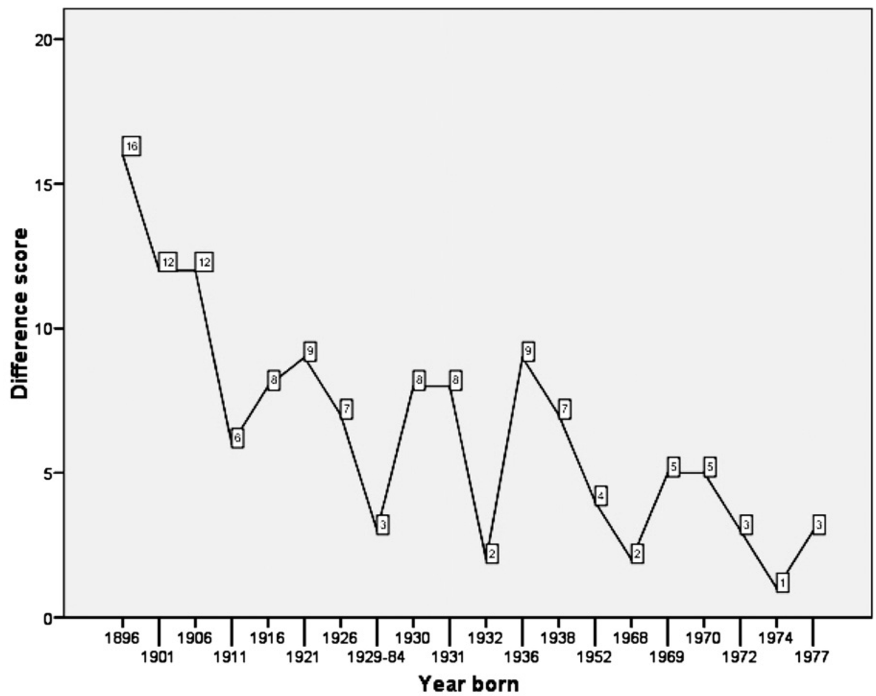

Fig. 2. Score differences between English-speaking and Afrikaans-speaking South Africans using all samples. Note. We used all data including data for which we only had information like standard deviations, here-say, and our own averaging out of data where no full scales were available; in other words everything in the Table above except three data-sets (the one outlier; the repeat data set of Biesheuvel and the Claassen et al. (1990) because the sample was born between 1929 and 1984 — leading to uncertainty as to which date to take.

Table 5
Comparable White groups taking the same test with several years in between by test battery (subtest level).

Test	Year born	Year sample	Gap	VIQ	NV IQ	Vgpd	NVgpd
NSAGT junior 1965 (van Eeden & Visser, 1992)	1971–1980	1987	22 years	104.85	107.75	2.2	3.52
NSAGT intermediate 1963 (van Eeden & Visser, 1992)	1971–1980	1987	24 years	106.19	110.62	2.57	4.42
NSAGT senior 1965 (van Eeden & Visser, 1992)	1971–1980	1987	22 years	102.6	107.54	1.18	3.42
NSAGT senior 1965 (Claassen, 1983)	1968	1981	16 years	96.76	100.07	−2.02	–

Table 6
Same groups taking different tests at a specific time (subtest level).

Test	VIQ	NVIQ	Gap	Vgpd	NVgpd
Sample of Whites tested in 1987 aged 7–16 born 1971–1980					
NSAGT junior 1965 and SSAIS-R (1987) (van Eeden & Visser, 1992)	104.85 and 101.78	107.75 and 100.78	22 years	1.39	3.16
NSAGT intermediate 1963 and SSAIS-R (1987) (van Eeden & Visser, 1992)	106.19 and 102.59	110.62 and 101.54	24 years	1.5	3.78
Sample of Whites tested in 1987 aged 6–14 born 1973–1981					
SSAIS (1962) and JSAIS (1976) (Robinson & Boshoff, 1990)	106.3 and 97.5	109.5 and 96.7	14 years	6.28	9.14
SSAIS (1962) and GSAT intermediate (1984) (Robinson & Boshoff, 1990)	106.3 and 92.3	109.5 and 94.5	22 years	6.36	6.81

a large gap the size of the GPD is very similar to the size for a gap of 17–29 years. The outcomes of the moderator type of test battery even show the opposite of what we expected. Table 7 shows a percentage of variance explained of 697%. This phenomenon is called “second-order sampling error”, and results from the sampling of studies in a meta-analysis. Percentages of variance explained greater than 100% are not uncommon when only a limited number of studies are included in an analysis. The proper conclusion is that all the variance is explained by statistical artifacts (see Hunter & Schmidt, 2004, pp. 399–401, for an extensive discussion).

We base our estimates of the White mean IQ on data from two different studies. Claassen et al. (2001) report standardization data for the South African WAIS-III based upon carefully sampled English-speaking South Africans from four groups: Blacks, Coloreds, Indians, and Whites. The South African version of the WAIS-III is virtually identical to the US version. The US norm tables were used and scaled scores for the total South African sample ($N=806$) were calculated after biased items had been replaced, yielding a mean Full Scale IQ of 94.79 (Table 9.2). The group of English-speaking Whites scores 7.78 IQ points above the mean of the total group (Table 9.5); correcting for the Flynn effect yields an IQ of 101.5 in comparison with the US mean.

Lynn and Vanhanen (2002, p. 219) cite a study by Owen (1992) using the Standard Progressive Matrices yielding a Flynn-effect-corrected IQ of 94 for Whites, compared with the British mean. This sample included both Afrikaans-speaking Whites and English-speaking Whites, so it should have a lower estimated value than the estimated value based on the WAIS-III, which used only English-speaking Whites. So, it appears that at present English-speaking Whites have a mean IQ highly

Table 7
Bare-bones meta-analytical results and moderator analyses.

Moderator	K	N	GPD	SD_{GPD}	% VE	80% CI
	17	6534	1.90	3.06	.2	−2.01–5.82
Method						
1 test 2 groups	4	3239	0.69	1.95	.2	−1.81–3.20
2 tests 1 group	13	3295	3.09	3.45	.3	−1.33–7.51
d (% Fluid and crystallized subtests)						
$d \geq 6\%$	8	1534	1.10	2.93	.3	−2.65–4.84
$-5\% \leq d \leq 5\%$	5	3476	1.26	2.81	.1	−2.35–4.86
$d \leq -6\%$	4	1524	4.19	2.56	.5	0.91–7.47
Gap (years)						
≤16	5	2796	1.63	4.02	.1	−3.51–6.77
17–29	9	3104	2.06	2.24	.4	−0.80–4.93
≥30	3	634	2.33	0	697	2.33–2.33
Mean age (years)						
>12	1	2053	−0.78			
10–12	15	4449	3.10	2.95	.3	−0.69–6.88
<10	1	32	8.07			
Quality sample						
Comparison with standardization sample	6	4289	1.12	1.87	.2	−1.26–3.52
Comparison not with standardization sample	11	2245	3.38	4.15	.3	−1.93–8.69

K = number of correlations; N = total sample size; GPD = mean observed Gain Per Decade (sample size weighted); SD_{GPD} = standard deviation of observed Gain Per Decade; %VE = percentage of variance accounted for by artifactual errors; 80% CI = 80% credibility interval.

similar to the mean IQs in Britain and the US, and Afrikaans-speaking Whites have a mean IQs that is a few points lower.

4. Conclusion

A Flynn effect could be identified for all our South African data based on no less than 41 combinations of datasets. There are clear gains per decade of about two-and-a-half IQ points for Whites and about one-and-a-half IQ point for Indians. In comparison to Whites in Europe and the United States, the Whites in South Africa show a somewhat smaller gain. As a group, White South Africans are quite westernized, so one could hypothesize that this explains a gain that comes close to that found in Western, industrialized countries. Greater gains are evidenced for non-verbal IQ scores as opposed to verbal IQ scores for the Whites, which is comparable to previous findings.

A special feature of the present paper is a comparison of test scores of Afrikaans- and English-speakers, starting with people born in 1896 and ending with people born in 1977. Over the course of approximately a century the large difference of about one *SD* in favor of English speakers diminishes by about three quarters. So, the group as a whole has a clear Flynn effect, but the effect is larger for the Afrikaans-speaking group. One could speculate that the diminishing gap between the Afrikaans- and English-speaking South Africans is driven partly by education and the diminishing gap in GDP between the two groups. However, there is no way to definitively prove this as trends in these two hypothesized causes and other hypothesized causes occur at the same time. There is some plausibility to the suggestion that highly-educated English-speaking Whites leaving the country in the last two decades is partly responsible for the narrowing of the gap between the two White groups.

Due to the questionable nature of some of the Black IQ datasets investigated but not used for this research (sample collection not always being explicitly stated) a major limitation of this paper is the lack of an estimate of the Flynn effect for the largest population group in the country.

There is a large variety of effects in the present study, so we tested for moderators using meta-regression. We tested for the moderators a) method used for computing the Flynn effect gain, b) type of test battery, c) time span, d) quality of the sample, and e) average age of sample, but found no clear support for moderators. The South African datasets should be included in a future meta-analysis; a larger number of data points allows more reliable tests of the moderators. Also, other moderators could be tested, such as time of the sampling.

References

Biesheuvel, S. (1952). Personnel selection tests for Africans. *South African Journal of Science*, *49*, 3–12.

Biesheuvel, S. (1952). The nation's intelligence and its measurement. *South African Journal of Science*, *49*, 120–138.

Biesheuvel, S., & Liddicoat, R. (1959). The effects of cultural factors in intelligence-test performance. *Journal of the National Institute for Personnel Research*, *8*, 3–14.

Bradmetz, J., & Mathy, F. (2006). An estimate of the Flynn effect: Changes in IQ and subtest gains of 10-yr-old French children between 1965 and 1988. *Psychological Reports*, *99*, 743–746.

Byrnes, R. M. (Ed.). (1996). *South Africa: A country study*. Washington: Library of Congress.

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, *27*, 703–722.

Claassen, N.C.W. (1983). The functioning of the NSAGT in various population groups. Unpublished MA thesis, University of Pretoria, South Africa.

Claassen, N. C. W. (1990). *Die meting van intelligensie in verskillende groepe met die Algemene Skolastiese Aanlegtoets (ASAT)* [The measurement of intelligence in different groups with the General Scholastic Ability Test (GSAT)]. South Africa: Pretoria: Human Sciences Research Council.

Claassen, N. C. W. (1997). Cultural differences, politics and test bias in South Africa. *European Review of Applied Psychology*, *47*, 297–307.

Claassen, N. C. W., De Beer, M., Hugo, H. L. E., & Meyer, H. M. (1991). *Manual for the General Scholastic Aptitude Test (GSAT) Senior Series.* South Africa, Pretoria: Human Sciences Research Council.

Claassen, N. C. W., Krynauw, A. H., Paterson, H., & wa ga Mathe, M. (2001). *A Standardisation of the WAIS-III for English-speaking South Africans.* South Africa, Pretoria: Human Sciences Research Council.

Claassen, N. C. W., Van Niekerk, H. A., & Kotzé, M. M. (1990). *Manual for the General Scholastic Aptitude Test (GSAT) Junior Series.* South Africa, Pretoria: Human Sciences Research Council.

Crawford-Nutt, D. H. (1976). Are Black scores on Ravens' Standard Progressive Matrices an artifact of method of test presentation? *Psychologia Africana*, *16*, 201–206.

Crawford-Nutt, D. H. (1977). The effect of educational level on the test scores of people in South Africa. *Psychologia Africana*, *17*, 49–59.

Cudeck, R., & Claassen, N. (1983). Structural equivalence of an intelligence test for two language groups. *South African Journal of Psychology*, *13*, 1–5.

Daley, T., Whaley, S., Sigman, M., Espinosa, M., & Neumann, C. (2003). IQ on the rise: The Flynn effect in rural Kenyan children. *Psychological Science*, *14*, 215–219.

Dent, G. R. (1949). *An investigation of certain aspects of Bantu intelligence.* South Africa, Pretoria: National Bureau of Educational and Social Research.

Department of Education (2008). *Education statistics in South Africa 2006.* South Africa, Pretoria: Department of Education.

Duckitt, J. (1992). Education and authoritarianism among English- and Afrikaans-speaking White South Africans. *The Journal of Social Psychology*, *132*, 701–708.

Fick, M. L. (1929). Intelligence test results of poor White, native (Zulu), Colored and Indian school children and the educational and social implications. *South African Journal of Science*, *26*, 904–920.

Fick, M. L. (1939). *An individual scale of general intelligence for South Africa.* South Africa, Pretoria: SA Council for Educational and Social Research.

Flynn, J. R. (1984). The mean of Q 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. (2006). Efeito Flynn: Repensando a inteligência e seus efeitos [The Flynn effect: Rethinking intelligence and what effects it]. In C. Flores-Mendoza, & R. Colom (Ed.), *Introdução a` Psicologia das Diferenças Individuais* (pp. 387–411). Brazil, Porto Alegre: ArtMed.

Flynn, J. R. (2007). *What is intelligence? Beyond the Flynn Effect.* Cambridge: Cambridge University Press.

Flynn, J. R. (2009). Requiem for nutrition as the cause of IQ gains: Raven's gains in Britain 1938–2008. *Economics and Human Biology*, *7*, 18–27.

Flynn, J. R. (2009). The WAIS-III and WAIS-IV: Daubert Motions favor the certainly false over the approximately true. *Applied Neuropsychology*, *16*, 98–104.

Foxcroft, C. D., & Aston, S. (2006). Critically examining language bias in the South African adaptation of the WAIS-III. *South African Journal of Psychology*, *32*, 97–102.

Furnham, A., Mkhize, N., & Mndaweni, T. (2004). Indian and isiZulu-speaking South African parents' estimates of their own and their children's intelligence. *South African Journal of Psychology*, *34*, 364–385.

Hunt, M. M. (1997). *How science takes stock: The story of meta-analysis.* New York: Sage.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Newbury Park, CA: Sage.

Huysamen, G. K. (1996). *Psychological measurement: An introduction with South African examples.* South Africa, Pretoria: van Schaik.

Irvine, S. H. (1969). Figural tests of reasoning in Africa: Studies in the use of Raven's Progressive Matrices across cultures. *International Journal of Psychology*, *4*, 217–228.

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

Kamin, L. J. (2006). African IQ and mental retardation. *South African Journal of Psychology*, *36*, 1–9.

Katz, Y. J. (1988). The relationship between intelligence and attitudes in a bilingual society: The case of White South Africa. *The Journal of Social Psychology*, *128*, 65–74.

Khaleefa, O., Abdelwahid, S. B., Abdulradi, F., & Lynn, R. (2008). The increase of intelligence in Sudan 1964–2006. *Personality and Individual Differences*, *45*, 412–413.

Knoetze, J., Bass, N., & Steele, G. (2005). The Raven's Colored Progressive Matrices : Pilot norms for isiXhosa-speaking primary school learners in peri-urban Eastern Cape. *South African Journal of Psychology*, *35*, 175–194.

Langenhoven, H. P. (1957). *Vergelyking van prestasies van Afrikaanse en Engelse groepe wat gelyk gemaak is ten opsigte van nie-verbale roupunttelling in the Nuwe Suid-Afrikaanse Groeptoets (NSAGT)*. [Comparison of performance of Afrikaans and English groups equated in terms of nonverbal raw score on the *New South African Group Test (NSAGT)*]. Pretoria: National Bureau for Educational and Social Research.

Liddicoat, R., & Roberts, A. O. H. (1962). Interim standardisation of the South African version of the Wchsler–Bellevue adult intelligence test. *Psychologia Africana*, *9*, 273–285.

Louw, P. E. (2004). *The rise, fall and legacy of Apartheid.* Westport CT: Praeger.

Luiz, D. M., & Heimes, L. (1994). A comparison between the junior South African Individual Scales (JSAIS) and the Griffiths Scales of Mental Development (Griffiths Scales). In R. van Eeden, M. Robinson, & A. B. Posthuma (Eds.), *Studies on South African Individual Intelligence Scales* (pp. 2–13). South Africa, Pretoria: Human Sciences Research Council.

Lynn, R. (2009). Fluid intelligence but not vocabulary has increased in Britain, 1979–2008. *Intelligence*, *37*, 249–255.

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

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

Mariotti, M. (2009). *Working papers in economics and econometrics.* Labour markets during Apartheid in South Africa. Working paper no 503. Canberra: Australian National University.

Meisenberg, G., Lawless, E., Lambert, E., & Newton, A. (2006). The social ecology of intelligence on a Caribbean Island. *Mankind Quarterly*, *46*, 395–433.

Morkel, P. (1950). *The A(G) mental alertness test: The reliability, difficulty, and grading of the forms*. Unpublished master's thesis, University of Stellenbosch, South Africa.

Must, O., te Nijenhuis, J., Must, A., & van Vianen, E. M. (2009). Comparability of IQ scores over time. *Intelligence*, *37*, 25–33.

Olckers, P. J. (1950). Skoolpunte en standaarde [School grades and standards]. *Journal for Social Research*, *1*, 172–182.

Owen, K. (1992). The suitability of Raven's Standard Progressive Matrices for various groups in South Africa. *Personality and Individual Differences*, *13* (2), 149–159.

Pressey, S. L., & Teter, G. F. (1919). A comparison of Colored and White children by means of a group scale of intelligence. *Journal of Applied Psychology*, *3*, 277–282.

Prinsloo, J. (2007). Making visible constructions of disadvantage through genealogical investigation: South African schooled literacies. *Critical Arts: A South–North Journal of Cultural and Media Studies*, *21*, 190–211.

Richter, L. M., Griesel, R. D., & Wortley, M. E. (1989). The Draw-a-Man test: A 50 year perspective on drawings done by black South African children. *South African Journal of Psychology*, *19*, 1–5.

Robinson, M., & Boshoff, J. S. (1990). *n' Vergelyking van psigometriese eienskappe van die JSAIS, die SSAIS en die OSAIS* [A comparison of psychometric properties of the JSAIS, the SAIS and the OSAIS]. South Africa: Pretoria: Human Sciences Research Council.

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. (2008). Testing the genetic hypothesis of group mean IQ differences in South Africa: Racial admixture and cross-situational consistency. *Personality and Individual Differences*, *44*, 768–776.

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

Schmidt, F. L., & Le, H. (2004). *Software for the Hunter–Schmidt meta-analysis methods [Computer software].* Iowa City: Department of Management and Organization, University of Iowa.

Shayer, M., Ginsburg, D., & Coe, R. (2007). Thirty years on — a large anti-Flynn effect? The Piagetian test Volume & Heaviness norms 1975–2003. *British Journal of Educational Psychology*, *77*, 25–71.

Shuttleworth-Edwards, A. B., Kemp, R. D., Rust, A. L., Muirhead, J. G. L., Hartman, N. P., & Radloff, S. E. (2004). Cross-cultural effects on IQ test performance: A review and preliminary normative indications on WAIS-III test performance. *Journal of Clinical and Experimental Neuropsychology*, *26*, 903–920.

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.

Smit, G. J. (1996). *Psychometrics: Aspects of measurement.* South Africa, Pretoria: Kagiso.

Sundet, J. M., Barlaug, D., & Torjussen, T. (2004). The end of the Flynn effect? A study trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence*, *32*, 349–362.

Teasdale, T., & Owen, D. (2008). Secular declines in cognitive test scores: A reversal of the Flynn Effect. *Intelligence*, *36*, 121–126.

van der Berg, A. R. (1989). Intelligence tests. In K. Oen, & J. J. Taljaard (Eds.), *Handbook for the use of psychological and scholastic tests of IPER and NIPR* (pp. 83–135). South Africa, Pretoria: Human Sciences Research Council.

van Eeden, R. (1991). *Manual for the Senior South African Individual Scale — Revised (SSAIS-R).* South Africa, Pretoria: Human Sciences Research Council.

van Eeden, R., de Beer, M., & Coetzee, C. H. (2001). Cognitive ability, learning potential, and personality traits as predictors of academic achievement by engineering and other science and technology students. *South African Journal of Higher Education*, *15*, 171–179.

van Eeden, R., & Visser, D. (1992). The validity of the Senior South African Individual Scale—Revised SSAIS-R) for different population groups. *South African Journal of Psychology*, *22*(3), 163–171.

Verster, J. M., & Prinsloo, R. J. (1988). The diminishing test performance gap between English speakers and Afrikaans speakers in South Africa. In S. H. Irvine, & J. W. Berry (Eds.), *Human abilities in cultural context* (pp. 534–559). Cambridge: Cambridge University Press.

Wicherts, J. M., Dolan, C. V., Carlson, J. S., & van der Maas, H. L. J. (2010). Raven's test performance of sub-Saharan Africans; mean level, psychometric properties, and the Flynn Effect. *Learning and Individual Differences*, *20*, 135–151.