Available online at www.sciencedirect.com
ScienceDirect
Intelligence 35 (2007) 1–12
INTELLIGENCE

General mental ability in South Asians: Data from three Roma (Gypsy) communities in Serbia

J. Philippe Rushton [a,*], Jelena Čvorović [b], Trudy Ann Bons [a]

[a] The University of Western Ontario, Canada
[b] Serbian Academy of Sciences, Belgrade, Serbia

Received 21 June 2006; received in revised form 5 September 2006; accepted 5 September 2006
Available online 18 October 2006

Abstract

To examine whether the Roma (Gypsy) population of Serbia, like other South Asian population groups, average lower than Europeans on *g*, the general factor of intelligence, we tested 323 16- to 66-year-olds (111 males; 212 females) in three different communities over a two-year-period on the Raven's Colored and/or Standard Progressive Matrices and four measures of executive function. Out of the total of 60 Matrices, the Roma solved an average of 29, placing them at the 3rd percentile on 1993 U.S. norms, yielding an IQ equivalent of 70. On the executive function tests, the Roma averaged at about the level of Serbian 10-year-olds. The Matrices showed a small mean sex difference favoring males. External validity was demonstrated by correlating the scores on Matrices with measures such as cranial capacity ($r=0.13$, $P<0.01$), spousal similarity ($r=0.17$, $P<0.05$), age at birth of first child ($r=0.26$, $P<0.01$), number of offspring ($r=-0.20$, $P<0.01$), and responsible social attitudes ($r=0.10$, $P<0.05$). Comparisons with extant data showed that items found difficult or easy by the Roma were those found difficult or easy by White, Indian, Colored, and Black South African 14- to 16-year-olds and by Black South African undergraduates ($rs=0.90$). There was no evidence of any idiosyncratic cultural effect. Instead, Roma/non-Roma differences were found to be most pronounced on *g*. This was shown by item-total correlations (estimates of the item's *g* loading), which predicted the magnitude of Roma/non-Roma differences on those same items, regardless of from which sample the item-total correlations were calculated, and by confirmatory factor analysis. The results indicate the remarkable cross-cultural generalizability of item properties across South Asians, Europeans, and sub-Saharan Africans and that these reflect *g* more than culturally specific ways of thinking.

Keywords: General factor; Cross-cultural; Progressive matrices

As the trend toward a more global economy continues, mean group differences in cognitive performance are likely to become more salient, both within and across countries. Most studies have been local in focus. In the United States they have been largely concerned with Whites, Blacks, Hispanics, East Asians and Native American Indians. In Australia they have been concerned with the lower mean scores of the Aborigines, and in New Zealand of the Maoris. Although a few theorists (e.g., Lynn & Vanhanen, 2002; Rushton & Jensen, 2005) have taken a global perspective and posited genetic and evolutionary explanations (50% genetic–50% cultural) for differences among the three

* Corresponding author. Department of Psychology, University of Western Ontario, London, Ontario, Canada N6A 5C2. Tel.: +1 519 661 3685.
E-mail address: Rushton@uwo.ca (J.P. Rushton).

0160-2896/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2006.09.002

macro-races of East Asians, Europeans and Africans, most hypotheses about such differences have focused on local cultural factors such as poverty and racism.

In the US, group inequalities were examined in Herrnstein and Murray's (1994) *The Bell Curve*, which re-analyzed data on 11,878 youths (3022 of who were African American) from the 12-year National Longitudinal Survey of Youth (NLSY). It found that most 17-year-olds with high scores on the Armed Forces Qualification Test, regardless of ethnicity, went on to occupational success by their late 20s and early 30s, while many of those with low scores went on to welfare dependence. It also showed that the average IQ for "African" Americans was lower than those for "Latino," "White," "East Asian," and "Jewish" Americans (IQs = 85, 89, 103, 106, and 115, respectively, pp. 273–278).

Internationally, Richard Lynn (2006) has gone beyond the traditional three macro-races and devoted a chapter to each of ten "genetic clusters" or population groups identified by Cavalli-Sforza, Menozzi, and Piazza (1994) in their mammoth *History and Geography of Human Genes*. Lynn tabulated 620 studies of IQ scores in 113 different countries from the beginning of the twentieth century to the present (N = 813,778) and found the world average IQ is 90 (Fig. 1). The East Asians (Chinese, Japanese and Koreans) obtained the highest mean IQ at 105. Europeans followed with an IQ of 100. Some ways below these were the Inuit or Eskimos (IQ 91), South East Asians (IQ 87), Native American Indians (IQ 87), Pacific Islanders (IQ 85), and South Asians and North Africans (IQ 84). Well below these were the average scores for the sub-Saharan Africans (IQ 67) followed by the Australian Aborigines (IQ 62). The lowest mean scores were obtained for the Bushmen of the Kalahari Desert and the Pygmies of the Congo rain forests (IQ 54).

The world's population groups are obviously not interchangeable. Some groups have proven so intractably below others in average test scores and concomitant differences in standards of living, educational outcome, and related phenomena that debates over potential remedial treatments have spanned generations. Regardless, long-standing group inequalities pose a problem in developing countries too such as India, Malaysia, Sri Lanka, Nigeria, and South Africa as well as the United States (Klitgaard, 1986; Sowell, 2004).

Lynn and Vanhanen (2002, 2006) found that average national IQ scores correlated 0.68 with per capita income and proposed that national IQs helped to explain why some countries are rich and others poor. This was a bold claim. Most economists regard it as axiomatic that all peoples of the world have the same IQ (e.g., Hanushek & Kimko, 2000). Lynn and Vanhanen showed that the evidence belies this assumption. They found national differences in intelligence that ranged between averages of 67 in sub-Saharan Africa to 105 in the "Asian tiger" economies of the Pacific Rim. They calculated the IQs for 113 countries and estimated them for others (e.g.,

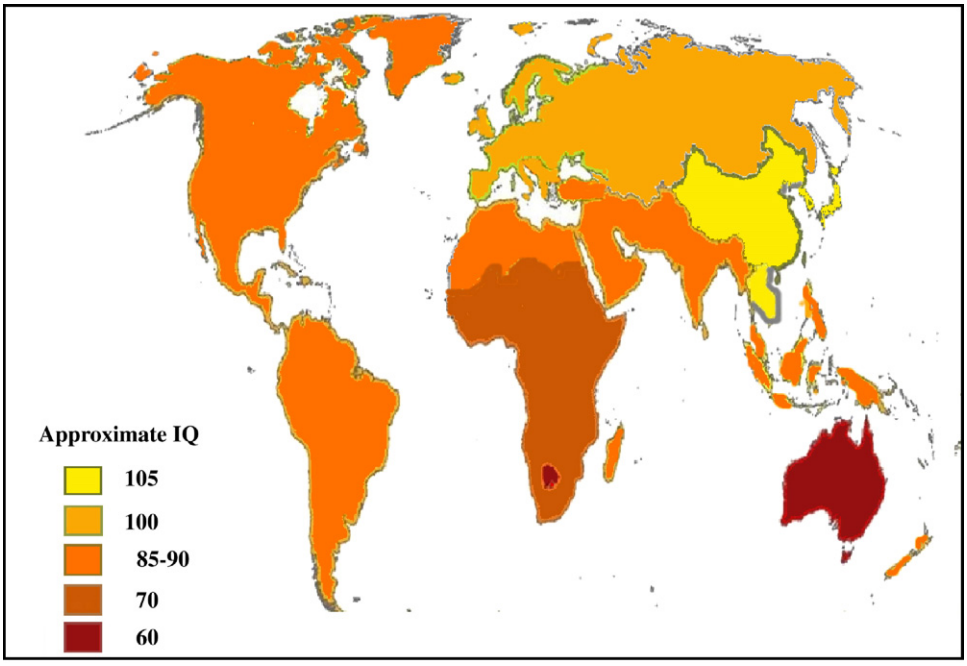

Fig. 1. World distribution of IQ scores of indigenous peoples (Adapted from Lynn, 2006).

Latvia at 98 from the measured IQ of 99 in Estonia and 97 in Russia; total $N = 185$).

To establish reliability, Lynn and Vanhanen (2006) showed that the correlation between independently measured general mental ability (GMA) test scores within countries was 0.95. To establish validity, they showed the national IQs correlated with the same nation's scores on tests of mathematics and science $r = 0.79$ to 0.89 (see also Lynn & Mikk, in press). Separately, Lynn (2006) also addressed construct validity more fundamentally, and concluded that the IQs of the nations could be explained in terms of the racial composition of the populations. Thus, the 6 East Asian nations (China, Japan, South Korea, Taiwan, Hong Kong and Singapore) all average IQs in the range of 105 to 108; the 30 European nations all average IQs in the range of 92 to 107; while the 19 nations of sub-Saharan Africa all average IQs in the range of 59 to 73.

Lynn's work has received a mixed reception. While some have usefully extended its analyses (e.g., Dickerson, 2006; Templer & Arikawa, 2006; Whetzel & McDaniel, 2006), others have criticized its reliance on inappropriate methods for calculating IQs and on poor sampling of populations and tests. Ervik (2003) wrote, "The authors fail to present convincing evidence and appear to jump to conclusions" (p. 522). Barnett and Williams (2004) asserted that the national IQs were "virtually meaningless," while Hunt and Sternberg (2006) said they were "technically inadequate" (p. 136).

A potentially contentious facet of the global IQ debate is the "under-achievement of South Asians" — the people from Bangladesh, India, Pakistan, Iraq, Iran, the Gulf States, the Near East, Turkey, and North Africa. Many of these countries are Muslim and are becoming increasingly associated by some notable commentators with a "Clash of Civilizations" (Huntington, 1998) as well as the longer standing "North–South divide." South Asians are closely related to Europeans and in some taxonomies of classical anthropology the two peoples have been regarded as a single race designated as the Caucasoid (e.g., Coon, Garn, & Birdsell, 1950; Rushton, 1995). In more recent analyses South Asians and North Africans emerge as a distinctive genetic "cluster" that differentiates them from Europeans (Cavalli-Sforza et al., 1994).

Lynn (2006) reviewed 37 IQ studies of South Asians from Sri Lanka, Nepal, India, Pakistan, Turkey, Iran, and Iraq and found an IQ range of from 77 to 96 with a median (and mean) of 84. He also reviewed 13 studies of recent immigrants from those countries in the UK and Australia and found an IQ range of from 83 to 97 with a median of 89 (mean = 90), with slightly higher scores for the second generation. He reviewed 18 further studies of South Asians and North Africans in Continental Europe, mainly in the Netherlands, and found an IQ range of from 75 to 94 with a median of 84 (mean = 85). Similar results in the Netherlands have been found by te Nijenhuis, Tolboom, Resing, and Bleichrodt (2004). Lynn reviewed 9 further studies of Indians in Africa, Fiji, Malaysia, and Mauritius and found an IQ range of from 77 to 91, with a median of 88 (mean = 86). Finally, Lynn reviewed 13 studies of South Asian and North African high school and university students and found an IQ range of from 85 to 106 with a median of 92 (mean = 94), eight points higher than that of general population samples.

Although the results of the 90 studies reviewed by Lynn (2006) were highly consistent in showing a 15 IQ point difference between Europeans and South Asians, which Lynn interpreted in terms of the general mental ability factor (*g*), none of the studies explicitly tested whether the differences were in fact on *g*. Nell (2000) elaborated the main alternative hypothesis to *g* for why non-Western populations score lower than Westerners; namely, that cultural differences imbue test items with different meanings and that non-Western groups are less test-wise, less interested, more anxious, work less efficiently, or give up sooner on items they find difficult. Given the concerns raised by Hunt and Sternberg (2006) that it is naive to ignore the possible social and political ramifications of research on global IQ, its relation to skin color, and the assertions that may be made about such findings in the secondary literature, more research is clearly called for.

In the present paper we provide a highly particular test of a very general prediction from Lynn's (2006) work on global IQ—specifically, that adult Gypsies (or Roma as they are now often called), a diverse population of South Asian stock who migrated to Eastern Europe from northwest India between the 9th and 14th centuries, have scores similar to their South Asian counterparts from whom they have been separated for centuries, and that the differences are more on the *g* loaded than culturally influenced items. The Roma have had a very different history in the intervening period than other South Asians. They retain a brown-skinned, East Indian appearance, and their geographic origin has been confirmed by linguistic analysis of their Romani language (Pearson, 1985) as well as by genetic sequencing studies (Fraser, 1995; Gresham et al., 2001). For the most part they have not intermarried with native Europeans and have retained their cultural traditions. Current estimates of the total Roma population size in Egypt range from 4 to 10 million, with the largest numbers concentrated in central and southeastern Europe (Marushiakova & Popov, 2001).

Although previous studies of Roma IQ have been conducted — all on children, mostly on small samples of unknown representativeness — they do consistently show an IQ range of from 70 to 83 (Bakalar, 2004; Raven, Court, & Raven, 1995; Save the Children, 2001). For example, Bakalar (2004) reviewed 10 studies in the Czech Republic and Slovakia with sample sizes that ranged from 33 to 178, with ages from 3- to 18-years, on both verbal and non-verbal tests, and with an IQ range of from 71 to 82 (median = 75; mean = 76). The most comprehensive of these used the Czech version of the Wechsler Intelligence Scale for Children-III (WISC-III) on a representative sample of 6- to 17-year-olds in the Czech Republic. The 89 Roma children averaged an IQ of 80 (Verbal IQ = 82; Performance IQ = 80) and the 1357 non-Roma children, an IQ of 101 (Verbal IQ = 101, Performance IQ = 101). Bakalar also found that Gypsy educational achievement was commensurate, with 62% of Roma children attending special education schools in comparison to 4% of the general population. However, Bakalar missed the largest study of Roma in Slovakia (N = 728), which was reported in the Raven's Manual for the Colored Progressive Matrices. It showed an IQ equivalent for 5- to 8-year-olds of 83 (Raven et al., 1995).

Čvorović (2004) found similar results in the former Yugoslavia, where illiteracy among Roma is a major problem. According to the 1981 census, the number of Roma without elementary school education was 80%; only 4% finished high school, and 0.2% college or university. According to Save the Children (2001), the situation had not changed much in the interim: 62% of Roma had not completed primary education. It reported that school psychologists who administered the Wechsler Intelligence Scale for Children-Revised (WISC-R) found that Roma children averaged an IQ of about 70, which placed many in the retarded category (p. 164). Since they were not able to handle the curriculum, they were sent to special schools. Serbian psychologists consider the Roma population to have a high percentage of "pseudo-retarded children" — children who score below the normal range on IQ tests but who are functionally normal in other ways. There are no overall data on the percentage of Roma attending schools for children with mental disabilities, although the data show that 70 to 80% of the children attending such schools are Roma (Save the Children, 2001).

Some observers attribute the poor education and low IQ scores of Roma to language barriers because many speak Romani, and to Roma culture, which has never produced a literary tradition. One survey revealed that 36% of the parents wanted their children to finish only 4 of the 8 obligatory grades of elementary school and that 18% were undecided on whether or not they wanted *any* education for their children (Save the Children, 2001). Upward mobility is not seen as a priority for Roma and may be considered harmful if it takes children from the community. A large number of Roma girls marry at 13- to 16-years. Many Roma parents register their children in school only when they need to collect welfare or other forms of social assistance, and continue to send their children to school only if provided a free meal. Roma parents are often pleased to accept the evaluation of their child as mentally disabled because it allows them to access various benefits such as free meals, medical care, and other humanitarian aid (Save the Children, 2001). It should be noted that in Eastern Europe education is free from preschool through to university.

However, even for those children who speak Serbian as a mother tongue, the situation is similar, with most children falling rapidly behind in school. A neuropsychiatrist working with Roma children on a daily basis in the Mental Health Institute of the Novi Beograd Medical Centre in Belgrade concluded that in general, Roma children, "don't know the language, and score poorly on tests. These children not only don't know Serbian, they don't know their own language either. Their parents are usually illiterate and have absolutely no appreciation of education" (Save the Children, 2001, p. 164).

1. Method

1.1. Overview

The primary purpose of this study is to examine general mental ability in three Roma communities in Serbia comprising a total of 323 adults ranging in age from 16 to 66 years. There was an excess of females (34% male; 66% female), which is typical of volunteer samples — as happens routinely, for example, when recruiting twins (Lykken, Tellegen, & DeRubis, 1978). The communities (i.e., Drenovac, Mirijevo, and Rakovica) are in the vicinity of Belgrade and were previously studied by the second author who established an excellent working relationship with them (Čvorović, 2004). The Roma were not paid for their participation, but small gifts were given, such as coffee for the adults and candy for the children.

Drenovac, a community in western Serbia (population = 2446), provided 100 of the Roma testees (35 male, 65 female). They belong to the Roma groups who entered Serbia from Romania in the 18th and 19th centuries. The village consists of 900 households, 10% of which have Roma inhabitants. Homes in the village have electricity and running water. The Roma speak

Serbian as their mother tongue and Romanian as a second language, and do not speak Romani at all. Historically, they tended to be musicians. Since the 1990s, males mainly work in factories while females stay home and raise families. Drenovac Roma consider themselves superior to other Roma communities and neither intermarry nor have social attachments with them. Their demography is similar to other Serbs, with an average age of first intercourse for females at 18 years, an average age of 22 at birth of first child, and an average of 2.2 children per family.

The Roma community of Mirijevo provided 89 testees (30 males, 59 females). This is a "typical" poor Roma city settlement. It is located in Belgrade, has about 250 houses, with 500 to 600 inhabitants, their number fluctuating with the season. The community consists of 3 groups: "Romanian" Gypsies who are Christian Orthodox, Ashkali who are Muslims, and Gurbeti who are for the most part Muslims but with some recent converts to Christianity. The settlement is poor: most live from social welfare/child allowances and occasional semi-skilled work. Only a few have full time jobs, mostly with the city garbage and waste collection company. Electricity and water exist as well as a partial sewage system.

The Roma community of Rakovica provided 134 testees (46 males, 88 females). The community is a semi-rural settlement, 7 km from Belgrade, inhabited by around 400 Roma from three different groups (Gurbeti, Ashkali and Romanian Gypsies). It is a relatively poor settlement although there is a partial sewage system and water and electricity. Only a dozen or so Roma are employed full-time. The inhabitants derive their income from social welfare/child allowances and occasional "private" business (the gathering of old newspapers, iron and black-market dealings).

1.2. Measures

The Raven's Standard Progressive Matrices (SPM) test was individually administered. This is probably the most well-known, most researched, and most widely used of all culture-reduced tests. Its popularity is evident from the fact that it has been used in well over 1000 studies (Raven, Raven, & Court, 1998). As an untimed "capacity" test, and even as a 20-minute "speed" or "efficiency" test, the results demonstrate reliability and validity across a wide range of populations. Retest reliabilities with an interval of approximately 1 year between administrations are between 0.83 and 0.93. Internal consistency coefficients of 0.80 are found across a wide variety of cultural groups (Raven et al., 1998). The Raven's was designed to measure Spearman's (1927) *g*, the general factor of intelligence, or at least the non-verbal component, and appears to be a good measure thereof (Jensen, 1998, p. 38). It is also described as a measure of "the ability to identify relationships," "analogical thinking," and the ability to "think clearly" (Raven et al., 1998).

The Standard Matrices consists of 60 diagrammatic puzzles, each with a missing part which the test taker attempts to identify from several options. The 60 puzzles are divided into five sets (A, B, C, D, and E) of 12 items each. In each set the first problem is as nearly self-evident as possible. The problems which follow build on the same reasoning as those that have gone before and provide opportunities to grasp mental operations required to solve the problems, which become progressively more difficult. To ensure sustained interest and freedom from fatigue, each problem is boldly presented, accurately drawn, and, as far as possible, pleasing to look at. No time limit is set and all testees are allowed to complete the test.

For a subset of the sample, we used the Colored Progressive Matrices (CPM), an "easier" test that has the effect of spreading out the scores of the bottom 20 percent of the general population. The CPM is typically given to young children, mentally impaired adolescents, and the elderly. It includes the two easiest sets from the SPM, plus a dozen additional items of similar difficulty and is made more attractive and less difficult by being in color.

Questionnaires also assessed Roma life history variables such as age of birth of first child, total number of children, and responsible social attitudes (Figueredo et al., 2005). In addition spouses were identified. These data form the basis of a separate paper (Figueredo et al., in preparation). Calipers were used to measure the external size of the head (length, width, and height) using standard procedures (Byers, 2002), and cranial capacity was calculated according to Lee and Pearson's (1901) formula as described in Rushton (1992).

For males:

$$\text{Cranial capacity } (\text{cm}^3) = 0.000337(L-11\text{ mm})(W-11\text{ mm})(H-11\text{ mm}) + 406.01$$

For females:

$$\text{Cranial capacity } (\text{cm}^3) = 0.0004(L-11\text{ mm})(W-11\text{ mm})(H-11\text{ mm}) + 206.6$$

where L, W, and H are length, width, and height in millimeters, with 11 mm subtracted to compensate for the fat and skin around the skull.

1.3. Procedure

The first round of testing took place throughout the 2004–2005 academic year (September to June). It was conducted by the second author in the test-takers' own home (or that of a neighbor) at a time free from distraction. As much as possible, the tests were made to seem like fun. The examiner sat near to the testees throughout, coaxed them, and told them not to worry about any difficulties. The subjects were given as long as they liked to answer each item. The goal throughout was to maximize the test-taker's score. Most Roma found the tasks very difficult; some complained of getting a "headache." They typically asked to stop the test before 30 min. After completing and analyzing 231 sets of scores on the SPM, it was decided to switch to the CPM. The remaining 92 subjects were administered the CPM. Although test-takers seemed to enjoy this version more, they continued to report the task was difficult and gave them a headache.

Since the Roma averaged an even lower score than expected – an IQ of 70 on both the Colored and the Standard versions of the test (see Results) – it was decided to carry out an additional assessment the following academic year (2005–2006) using four "executive function" type tests typically given to either preschool children or to neurologically impaired adults when a test of general mental ability is considered too difficult. The executive function tests were developed by the Institute for Experimental Phonetics and Speech Pathology in Belgrade (Baterija Testova, 2002) and were said to measure the ability to keep track of what has already been done, planning, co-ordination, as well as simple concepts, memory, independence, and ability to concentrate. These kinds of tests have also been recommended for cross-cultural research (Nell, 2000). We used the four tests described below.

Noun Definition test. Test-takers are asked, "What is a *man*?" (followed by: *mother*, *life*, *house*, and *sun*). Answers are scored from 1 = *low* to 8 = *high* on the specificity or generalizability of the answer based on a grading system (e.g., 3 = *answer based on function*).

Opposite Attributes test. Test-takers are given 4 words (*big*, *good*, *black*, and *free*) and asked to generate some of their opposites. Each item is scored 1 = *low* through 5 = *high*.

Verbal Memory test. This test consists of 8 sections comprising a total of 65 short sentences. Each response is scored 0 = *incorrect* or 1 = *correct*.

Drawing test. This test consists of a sheet of paper with a rectangle on it and the test taker is asked to draw a circle, a cross, and a triangle on top of the rectangle in a particular order and of about the same size. Each drawing is scored on 5 elements from 1 = *low* to 3 = *high* for a maximum score of 15.

2. Results

The Roma averaged very low scores on all tests, equivalent to an IQ of about 70. On both CPM and SPM, each item was scored as 0 (*incorrect*) or 1 (*correct*). The CPM scores were equated to SPM scores using the Manual's conversion table (Raven et al., 1998, Table SPM 4). The internal consistency coefficient (Cronbach's alpha) was 0.91 (after eliminating people with zero scores on at least 80% of the items, which otherwise would have inflated the alpha to 0.94). The Roma scored a total of 29 items correct out of 60, which is below the 3rd percentile for 18- to 22-year-olds on U.S. norms (Raven et al., 1998, Table SPM 13). These SPM percentile points convert to an IQ equivalent of about 70 (Raven et al., 1990, p. 98). The same IQ equivalent (70) was found when the analysis was limited to those doing only the CPM (N=92; CPM=25; equivalent to SPM=27).

On the executive function tests, the Roma also had very low scores. On the Noun Definitions test, out of a maximum score of 40, the Roma had a mean of 14, slightly lower than Serbian preschoolers with a score of 15. They gave either wrong or irrelevant answers, or answers from an immediate perspective, or that pointed only to the word's function such as "my mother is a kind woman," or "my house needs repair," or "the sun is a Gypsy stove," and they demonstrated little or no abstraction. On the Opposite Attributes test, out of a maximum of 20, the Roma and preschoolers had means of 16 and 20. On the Verbal Memory test, out of a possible 65, the Roma had a mean of 48; 42% achieved a perfect score on the first 5 sections of the test, the easiest part, compared to 95% of Serbian 6- to 7-year olds who did so. On the Drawing test, out of a maximum of 15, the Roma had a mean of 8, well below the level of Serbian 13-year-olds who achieve scores of between 12 and 15.

We examined the reliability of the Roma/non-Roma comparisons by seeing if they changed over the Roma age distribution. We compared the 16- to 22-year-olds (N=43) with the 23- to 66-year-olds (N=280). The younger age group tended to score the same as the older group, or marginally higher. When they did score higher they were still very low compared to Serbian children. On the Raven's SPM, the older and younger age group scored the same, as they did on the Opposite Attributes test, and on the Drawing test. On the Noun Definitions

test, the younger group scored higher than the older one (means = 15 and 14, respectively; $F(1321) = 4.90$, $P = < 0.05$), but only about the same as Serbian 1st graders. On the Verbal Memory test, 51% of the younger group earned a perfect score on the easiest section, compared to 40% of the older Roma, while 95% of Serbian 13-year-olds did so.

We examined the results separately for the three Roma communities (i.e., Drenovac, Mirijevo, and Rakovica). The Drenovac Roma scored higher than those in the other two communities. On the Raven's SPM, the Means were 35, 23, and 28 respectively (*SD*s = 11, 12, 10; $F(2,320) = 27.91$; $P < 0.01$). On the Noun Definitions test, the means were 16, 13, and 14, respectively (*SD*s = 3, 4, 3; $F(2,320) = 14.67$; $P < 0.01$). On the Opposite Attributes test, the means were 18, 15, and 14, respectively (*SD*s = 6, 8, 8; $F(2,320) = 7.54$; $P < 0.01$). On the Verbal Memory test, the means were 50, 47, and 48, respectively (*SD*s = 2, 5, 3; $F(2,320) = 13.77$; $P < 0.01$; with 58%, 34% and 35% getting perfect scores on the easiest section. On the Drawing test, the means were 10, 7, and 8, respectively (*SD*s = 3, 3, 4; $F(2,320) = 12.45$; $P < 0.01$).

We also examined sex differences. Males ($N = 111$) averaged higher than females ($N = 212$) on the Raven's SPM (means = 33, 27; *SD*s = 13, 11; $F(1321) = 22.29$; $P < 0.01$), yielding IQ equivalents of 74 and <70, but lower on the Opposite Attributes test (Means = 14, 16; *SD*s = 9, 7; $F(1321) = 7.87$; $P < .01$). There were no sex differences on the Noun Definition, the Verbal Memory, and the Drawing tests.

2.1. Item analyses and differences in g

Table 1 shows the proportion of Roma who selected the correct answer for each of the 60 items on the Standard Matrices ($N = 231$). Forty-one of the items (68%) had values within the optimal discrimination range of 0.20 to 0.80, with only 10 of the items (17%) proving "too difficult." For comparison purposes, Table 1 also shows the proportion correct from two South African studies. The first was by Rushton and Skuy (2000) of 173 Black African 17- to 23-year-old undergraduates at the University of the Witwatersrand (IQ = 85), and the second was by Owen (1992) of 1093 Black, 778 Colored, 1063 Indian, and 1056 White 14- to 16-year-olds (IQs given by Lynn [2006] as 74, 80, 91, and 94, respectively). For all groups, sets A and B were the easiest, followed by sets C and D, while set E was the most difficult. Across the 60 items, those found difficult by the Roma were found difficult by the five other groups (mean Pearson $r = 0.90$; $P < 0.01$; mean Spearman rho = 0.90, $P < 0.01$). This suggests the test measures the same construct in all six groups.

Table 1
Proportion of correct answers for each of the 60 items on the Raven's Standard Progressive Matrices for six culturally diverse samples

Item		Serbian Roma Adults	Black South African Undergraduates	Black South African 14- to 16-year-olds	Colored South African 14- to 16-year-olds	Indian South African 14- to 16-year-olds	White South African 14- to 16-year-olds
Sample size		231	173	1093	778	1063	1056
Set A	1	1.00	1.00	1.00	1.00	1.00	1.00
	2	.98	1.00	.99	.99	.99	.99
	3	.96	.99	.97	.98	.99	.99
	4	.92	.98	.96	.98	.99	.99
	5	.87	.97	.95	.98	.99	.99
	6	.89	.98	.95	.98	.99	.99
	7	.70	.94	.65	.87	.90	.98
	8	.62	.90	.83	.87	.91	.95
	9	.81	.96	.80	.93	.95	.99
	10	.69	.92	.67	.82	.87	.96
	11	.45	.82	.52	.75	.79	.90
	12	.34	.60	.33	.48	.53	.72
Set B	13	.90	.99	.96	.98	.99	.98
	14	.79	.98	.89	.96	.98	.99
	15	.78	.96	.75	.91	.95	.98
	16	.63	.90	.57	.87	.89	.97
	17	.63	.90	.45	.78	.84	.95
	18	.70	.86	.49	.72	.81	.88
	19	.43	.74	.45	.64	.74	.76
	20	.44	.77	.32	.61	.74	.81
	21	.51	.83	.40	.61	.79	.85
	22	.56	.85	.45	.74	.88	.95
	23	.30	.76	.32	.57	.76	.82
	24	.29	.56	.21	.41	.58	.64
Set C	25	.74	.94	.80	.91	.97	.97
	26	.77	.91	.72	.87	.94	.96
	27	.74	.92	.71	.85	.94	.95
	28	.58	.86	.51	.71	.80	.86
	29	.58	.88	.55	.84	.91	.93
	30	.40	.73	.46	.67	.74	.85
	31	.45	.86	.40	.67	.80	.89
	32	.25	.46	.29	.44	.50	.64
	33	.47	.72	.40	.65	.82	.78
	34	.22	.55	.20	.32	.45	.53
	35	.20	.46	.18	.22	.39	.42
	36	.07	.13	.07	.03	.05	.10
Set D	37	.81	.96	.78	.93	.96	.98
	38	.69	.94	.64	.86	.94	.96
	39	.57	.93	.58	.81	.90	.94
	40	.54	.89	.47	.76	.87	.89
	41	.58	.92	.61	.86	.93	.95
	42	.47	.87	.42	.72	.86	.90
	43	.33	.79	.35	.60	.70	.79
	44	.43	.80	.36	.60	.69	.77
	45	.33	.79	.31	.50	.62	.73
	46	.25	.73	.26	.47	.68	.78
	47	.23	.33	.11	.17	.24	.22
	48	.11	.21	.05	.09	.11	.15

(continued on next page)

Table 1 (*continued*)

Item		Serbian Roma Adults	Black South African Undergraduates	Black South African 14- to 16-year-olds	Colored South African 14- to 16-year-olds	Indian South African 14- to 16-year-olds	White South African 14- to 16-year-olds
Sample size		231	173	1093	778	1063	1056
Set E	49	.47	.77	.32	.52	.71	.79
	50	.39	.68	.22	.36	.66	.77
	51	.29	.64	.20	.43	.60	.67
	52	.24	.47	.13	.28	.46	.60
	53	.16	.50	.10	.25	.46	.64
	54	.17	.39	.11	.24	.36	.51
	55	.15	.38	.15	.23	.34	.38
	56	.14	.23	.07	.13	.25	.37
	57	.13	.25	.10	.14	.25	.34
	58	.04	.11	.04	.05	.10	.14
	59	.06	.06	.06	.05	.05	.06
	60	.06	.08	.08	.06	.09	.09

Another index of item similarity across groups is the item-total correlation, calculated using the biserial correlation of each item's pass or fail status (0 or 1) with the total score on the test (see Table 2). This statistic indicates the extent to which a particular item measures the same construct measured by the test as a whole, as well as how well the item discriminates between testees within each group. Across the 60 items, those items with high values among the Roma had high values in the other five groups (mean Pearson $r=0.40$; $P<0.01$; mean Spearman rho $=0.38$, $P<0.01$).

Since 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 (Table 2) also provides an estimate of each item's *g* loading. As a first test of whether Roma/non-Roma differences are more pronounced on the more *g* loaded items, we carried out Jensen's (1980) method of correlated vectors. To do this, we followed the procedure used by Rushton and Skuy (2000), Rushton, Skuy, and Fridjohn (2002), Rushton, Skuy, and Fridjohn (2003), Rushton, Skuy, and Bons (2004) and correlated the item-total correlations shown in Table 2 (which estimate *g*), with the standardized differences between Roma and non-Roma in proportion passing each item shown in Table 1 (which estimates the population-group effect size). The correlation between the *g* loadings and the absolute magnitude of the five sets of Roma/non-Roma differences averaged 0.54 ($P<0.01$; Pearson's *r*) and 0.53 ($P<0.01$, Spearman's rho) using the item-total correlations for the Roma group, and 0.59 ($P<0.01$; Pearson's *r*) and 0.55

Table 2
Item-total correlations for each of the 60 items on the Raven's Standard Progressive Matrices for six culturally diverse samples

Item		Serbian Roma Adults	Black South African Undergraduates	Black South African 14- to 16-year-olds	Colored South African 14- to 16-year-olds	Indian South African 14- to 16-year-olds	White South African 14- to 16-year-olds
Sample size		231	173	1093	778	1063	1056
Set A	1	.14	–	–	–	–	–
	2	.29	–	.13	.08	.06	.03
	3	.22	.11	.24	.21	.14	.09
	4	.31	.26	.24	.23	.09	.10
	5	.48	.32	.25	.28	.10	.09
	6	.47	.27	.29	.28	.17	.04
	7	.53	.44	.61	.49	.41	.14
	8	.54	.27	.28	.27	.31	.14
	9	.54	.44	.49	.38	.35	.20
	10	.48	.44	.54	.51	.41	.20
	11	.32	.51	.51	.47	.48	.20
	12	.48	.40	.45	.43	.42	.37
Set B	13	.41	.37	.23	.23	.21	.18
	14	.50	.46	.37	.28	.19	.23
	15	.52	.59	.63	.52	.48	.26
	16	.43	.50	.54	.54	.41	.28
	17	.52	.45	.56	.50	.50	.33
	18	.47	.40	.56	.44	.48	.25
	19	.43	.44	.50	.44	.37	.27
	20	.63	.51	.61	.48	.49	.36
	21	.59	.46	.60	.51	.47	.35
	22	.55	.56	.65	.59	.52	.40
	23	.59	.56	.58	.50	.50	.39
	24	.49	.51	.46	.43	.51	.43
Set C	25	.35	.48	.50	.40	.40	.33
	26	.37	.31	.49	.53	.39	.26
	27	.56	.57	.65	.47	.46	.31
	28	.50	.36	.54	.48	.37	.35
	29	.56	.51	.63	.57	.44	.37
	30	.63	.50	.56	.49	.45	.33
	31	.68	.52	.61	.53	.57	.41
	32	.44	.48	.46	.33	.41	.31
	33	.55	.50	.55	.44	.43	.31
	34	.52	.53	.40	.29	.40	.42
	35	.50	.40	.26	.28	.34	.30
	36	.03	.31	−.16	−.03	.13	.25
Set D	37	.54	.31	.61	.52	.40	.25
	38	.58	.51	.67	.59	.46	.35
	39	.52	.42	.62	.53	.41	.38
	40	.65	.40	.65	.58	.41	.39
	41	.64	.60	.67	.57	.43	.37
	42	.62	.42	.64	.62	.46	.36
	43	.58	.45	.58	.45	.41	.38
	44	.51	.54	.49	.42	.44	.37
	45	.54	.49	.49	.43	.46	.44
	46	.42	.57	.43	.43	.54	.54
	47	.44	.33	.12	.12	.20	.22
	48	.40	.36	.16	.11	.25	.23

Table 2 (*continued*)

Item		Serbian Roma Adults	Black South African Under-graduates	Black South African 14- to 16-year-olds	Colored South African 14- to 16-year-olds	Indian South African 14- to 16-year-olds	White South African 14- to 16-year-olds
Sample size		231	173	1093	778	1063	1056
Set E	49	.52	.38	.47	.47	.39	.36
	50	.60	.45	.36	.37	.45	.48
	51	.40	.50	.37	.36	.47	.47
	52	.53	.51	.27	.39	.58	.55
	53	.36	.54	.24	.40	.56	.55
	54	.44	.41	.25	.35	.42	.48
	55	.35	.33	.21	.19	.34	.35
	56	.38	.49	.10	.29	.45	.42
	57	.38	.35	.15	.25	.37	.41
	58	.05	.15	.04	.04	.24	.29
	59	−.14	.21	−.00	−.05	.06	.09
	60	.22	.48	−.11	−.01	.11	.16

($P<0.01$, Spearman's rho) using the item-total correlations for all the non-Roma groups. (*Note*: 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.)

To test further whether the Roma/non-Roma differences can be attributed in part to the *g* factor, we performed a multi-group confirmatory factor analysis (MGCFA) of the single-factor *g* solution. We used the AMOS 6.00 program (Arbuckle, 2005) to compare scores from the Roma ($N=231$) with those from the African undergraduates ($N=173$) for whom we also had item level raw data (Rushton & Skuy, 2000). We used an iterative procedure to create six parcels of 10 items each matched for difficulty level to normalize the distributions (Fig. 2). A Chi-Square test was conducted to determine if the *g* loadings for both groups were the same. To do this, the model parameters were first estimated with the loadings for all groups constrained to the same value, and then with the constraint removed, thereby letting each group's factor loading have a different value. If the Chi-Square of the difference between the constrained and unconstrained models is not significant, the loadings on *g* are the same for both groups; if the Chi Square is significant, separate parameter estimates for each group would provide a better fit. We used a maximum likelihood procedure and the chi-square value is based on the model with loadings constrained across groups. The results indicated that the same model-single-factor *g*-fit the data for both groups (e.g., $\Delta\chi^2=4.10$, df=5, ns; CFI=0.99; TLI=0.99; IFI=0.99; GFI=0.96; also RMSEA=0.05, $\chi^2=43.43$, df=23,

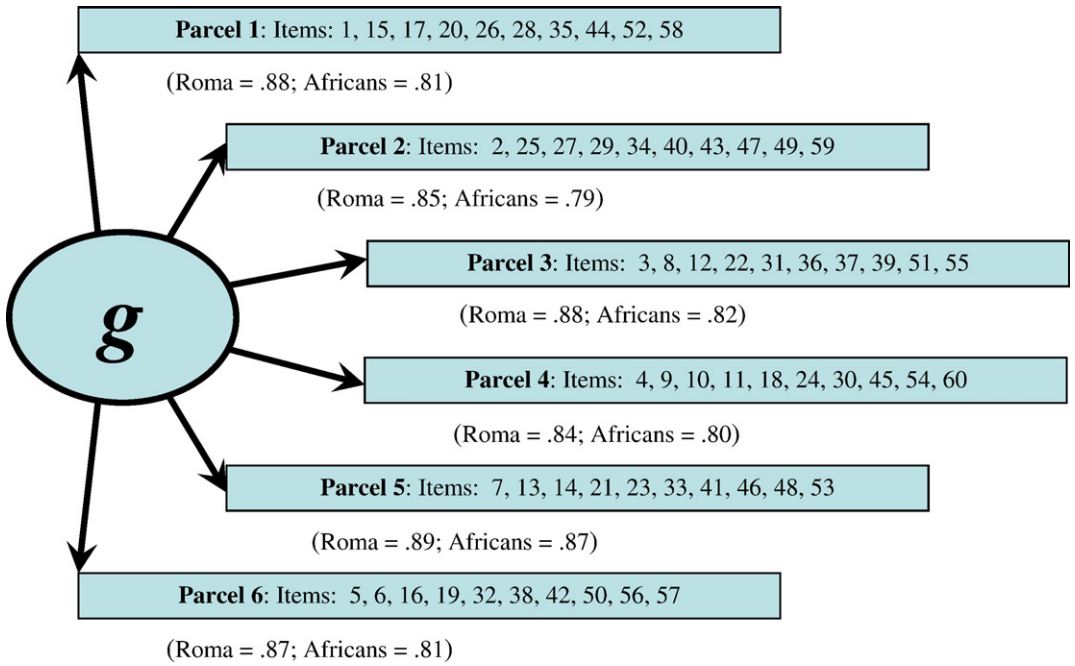

Fig. 2. Factor structure model of *g* loadings for 6 parcels of items matched for difficulty level to test goodness-of-fit of *g* factor between Serbian Roma Adults ($N=231$) and South African Black undergraduates ($N=173$). Also shown are the factor loadings on *g* of each item parcel, separately for the Roma and the Africans.

$P<0.01$). The reliable variance in these 10 parcels due to *g* is: for the Roma, 79%; and for the African undergraduates, 72%. The standardized factor solution for the hypothesized model is shown in Table 3.

We also tested the single-factor *g* model in other ways. For example, we parceled the items based on the three lower order factors that Lynn, Allik, and Irwing (2004) extracted from the Raven's administered to 2735 Estonian adolescents, viz, *gestalt continuation*, *verbal-analytical reasoning*, and *visuospatial ability*. Good fit indices were obtained although the model was only marginally significant ($\chi^2=32.49$, df,=23, $P<0.10$). Since our samples were small and many of the items had non-optimal pass rates, we also tested the model selecting only those 19 items in Table 1 with pass rates between 20 and 80 percent for both groups. The single-factor *g* model was confirmed at the item level for both groups, although with fit indicators of only 0.85. When we aggregated the 19 items into parcels three at a time the model fits rose to 0.95. In short, the factor structure of cognitive ability seems to be equivalent for both the Roma and the African undergraduates.

2.2. Construct validities

We examined the construct validity of the scores on the Raven's Matrices by correlating the measures with cranial capacity ($r=0.13$, $P<0.01$), spousal similarity ($r=0.17$, $P<0.05$), age at birth of first child ($r=0.26$, $P<0.01$), number of offspring ($r=-0.20$, $P<0.01$), and responsible social attitudes ($r=0.10$; $P<0.05$). All these correlations, while low, are in the expected direction and show the test has validity. On the other hand, while the four executive function tests correlated with each other (mean $r=0.38$; $P<0.01$), they failed to correlate with the Raven's ($r=0.04$) or for the most part the 5 external criteria of cranial capacity ($r=-0.03$; *ns*), spousal similarity ($r=0.01$; ns), age at birth of first child ($r=0.14$; $P<0.05$), total number of offspring ($r=-0.08$, ns), and social attitudes ($r=0.04$, ns). A multiple regression on the five external criteria found: for Raven's, $R=0.41$; for executive function tests, $R=0.19$.

Table 3
Standardized factor solution for hypothesized *g* factor model

Item parcel number	Serbian Roma	African undergraduates
1	.881	.810
2	.849	.790
3	.882	.821
4	.836	.797
5	.887	.874
6	.869	.808

3. Discussion

This is the first study of which we are aware to provide a comprehensive estimate of the IQ of an *adult* Roma. We found an IQ of 70 on both the Raven's Standard and the Raven's Colored Progressive Matrices, as well as on four "executive function" tests. The Roma sample was large ($N=323$) and the results held up over a two year test period, for both younger and older age groups, for both sexes, across three communities including Drenovac where the Roma speak the majority language and are adapted to Serbian cultural norms, as well as across several comparison groups such as Serbian preschoolers, White, Indian, Colored, and Black South African high school students, and Black South African undergraduates.

The results reported here conform to those from the 100 studies of South Asians reviewed in the Introduction, which showed an IQ range of from 70 to 106 (median=85; mean=86). Even a few years ago, reporting a mean IQ of 70 for any population group would have been considered not only an absurdity but also an injustice (Nell, 2000). Yet new empirical work continues to accumulate finding that the world mean IQ is 90, that mean IQs of 70 are found routinely in sub-Saharan Africa, and that mean IQs of 70 to 90 are typical of many other regions of the world (e.g., Lynn, 2006; Rushton et al., 2004; Sternberg et al., 2001, 2002). Mean IQs as high as 100 are seldom found outside of European and East Asian population groups (see Fig. 1).

This paper does not address the question of what mix of genetic and environmental factors are involved in the mean population group differences, and a wide range of interpretations is possible. The scores of the Roma reported for this sample might be largely cultural in origin and may seriously underestimate the true Roma population mean. It is a truism that test takers should be similar in cultural, educational, and social background to those on whom the test was standardized. If a group differs markedly from the standardization sample, the use of the norms may be inappropriate. The social context of the Roma in Serbia differs significantly from the majority Serbian population. Roma children grow up in culturally disadvantaged conditions, often live in overcrowded homes, and are not as exposed to the intellectual stimulation and test taking attitudes typically associated with high test scores. There is also much evidence that a sub-optimal level of nutrition has an adverse effect on general intelligence. Another

possibility too is that the higher scoring Roma have moved elsewhere in Europe — many of the individuals in the Mirijevo community reported having relatives in Germany.

Questions can be raised about both the internal and external validity of the tests used to assess the Roma in this study, as well as the adequacy of the sampling procedures, and more data would be useful. With respect to internal validity on the SPM, those items found easy or difficult by the Roma were the same ones found easy or difficult by the non-Roma. The estimates of *g* from these items predicted the magnitude of the differences between the Roma and several comparison groups, suggesting that the test measured the same construct in all groups. Nonetheless, criticisms have been made of one of our procedures — Jensen's (1998) method of correlated vectors (Ashton & Lee, 2005; Dolan, Roorda, & Wichert, 2004), although we conceptually replicated the results using confirmatory group factor analysis. Also, the Raven's correlated with a number of other external criteria such as cranial capacity ($r=13$; $p<0.01$), spousal similarity ($r=0.17$; $p<0.05$), age at birth of first child ($r=0.26$; $p<0.01$), number of offspring ($r=-0.20$; $p<0.01$), and responsible social attitudes ($r=0.10$; $p<0.05$). However, it did not correlate with the four executive function tests ($r=0.04$).

The failure of the Raven's test to correlate with the executive function tests led one reviewer to question whether the Raven's was a measure of *g* in this sample because any *g* loaded test must, by definition, predict performance on other tests of intelligence. On the other hand, a failure to find a predicted correlation can be due as much to the weakness of the criterion as of the predictor variable (Jensen, 1980, p. 310). This seemed to be the case here because, although the four executive function tests correlated with each other (mean $r=.38$), as we showed, they failed to correlate with the other criteria. There is also some evidence that executive function tests are only loosely related to working memory capacity, and hence to *g* (Oberauer, Suss, Wilhelm, & Wittman, 2003). If so, this would solve the apparent anomaly of the executive function tests not correlating with *g*. Nonetheless, given that weaknesses are inevitably to be found in any one study regarding the sampling, the assessments, and the statistical analyses, we call for more hypothesis tests of whether *g* underlies these and other population group differences (e.g., Hartmann et al., 2007).

Nell (2000) elaborated the main alternative hypothesis to *g* for why non-Western populations score lower than Westerners; namely, that cultural differences imbue test items with different meanings and that non-Western groups are less test-wise, less interested, more anxious, work less efficiently, or give up sooner on items they find difficult. The degree of cross-situational generalizabilty found in the Raven's Matrices, however, makes it much less likely that the differences could be caused by biased test instruments or culturally idiosyncratic ways of answering questions. Psychometricians might accept, therefore, that within limits, the SPM test scores provide a useful index of current cognitive functioning. The burden of proof has shifted to the critics to provide new datasets or superior methods of analysis, rather than mere verbal arguments, if they wish to maintain that such test scores are neither reliable nor valid.

References

Arbuckle, J. L. (2005). *AMOS 6.00 user's guide.*Chicago: SPSS Inc. Computer software.

Ashton, M. C., & Lee, K. (2005). Problems with the method of correlated vectors. *Intelligence*, *33*, 431−444.

Bakalar, P. (2004). The IQ of Gypsies in central Europe. *Mankind Quarterly*, *44*, 291−300.

Barnett, S. M., & Williams, W. (2004). IQ and the wealth of nations: Review. *Contemporary Psychology*, *49*, 389−396.

Baterija Testova (2002). *Institut za Eksperimentalnu Fonetiku i Patologiju Govora, Beograd, Serbia. [Test Battery: Institute for Experimental Phonetics and Speech Pathology, Belgrade, Serbia.].*

Byers, S. N. (2002). *Introduction to forensic anthropology: A textbook.* Boston, MA: Allyn & Bacon.

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

Coon, C. S., Garn, S. M., & Birdsell, J. B. (1950). *Races.* Springfield, IL: Thomas.

Čvorović, J. (2004). Sexual and reproductive strategies among Serbian Gypsies. *Population and Environment*, *25*, 217−242.

Dickerson, R. E. (2006). Exponential correlation of IQ and the wealth of nations. *Intelligence*, *34*, 291−295.

Dolan, C. V., Roorda, W., & Wicherts, J. M. (2004). Two failures of Spearman's hypothesis: The GATB in Holland and the JAT in South Africa. *Intelligence*, *32*, 155−173.

Ervik, A. O. (2003). Review of IQ and the wealth of nations. *The Economic Journal*, *113*, F406−F408.

Figueredo, A. J., Cvorovic, J., et al. (in preparation). *Assortative mating and life history variables in cross-cultural perspective.* Manuscript, Department of Psychology, University of Arizona, Tucson, Arizona, 85721.

Figueredo, A. J., Sefcek, J. A., Vasquez, G., Brumbach, B. H., King, J. E., & Jacobs, W. J. (2005). Evolutionary personality psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology,* (pp. 851−877). New York: Wiley.

Fraser, A. (1995). *The Gypsies.* Cambridge, MA: Blackwell.

Gresham, D., Morar, B., Underhill, P. A., Passarino, G., Lin, A. A., Wise, C., et al. (2001). Origins and divergence of the Roma (Gypsies). *American Journal of Human Genetics*, *69*, 1314−1331.

Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor force quality, and the growth of nations. *American Economic Review*, *90*, 1184−1208.

Hartmann, P., Krusse, N. H. S., & Nyborg, H. (2007). Testing the cross-racial generality of Spearman's hypothesis in two samples. *Intelligence*, *34*, 47−57.

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

Hunt, E., & Sternberg, R. (2006). Sorry, wrong numbers: An analysis of a study of a correlation between skin color and IQ. *Intelligence*, *34*, 131−137.

Huntington, S. P. (1998). *The clash of civilizations and the remaking of world order.* New York: Simon & Schuster.

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

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

Klitgaard, R. (1986). *Elitism and meritocracy in developing countries: Selection policies for higher education.* Baltimore, MD: Johns Hopkins University Press.

Lee, A., & Pearson, K. (1901). Data for the problem of evolution in man: VI. A first study of the correlation of the human skull. *Philosophical Transactions of the Royal Society of London*, *196A*, 225−264.

Lykken, D. T., Tellegen, A., & DeRubis, R. (1978). Volunteer bias in twin research: the rule of two thirds. *Social Biology*, *25*, 1−9.

Lynn, R. (2006). *Race differences in intelligence: An evolutionary analysis.* Augusta, GA: Washington Summit Books.

Lynn, R., Allik, J., & Irwing, P. (2004). Sex differences on three factors identified in Raven's Standard Progressive Matrices. *Intelligence*, *32*, 411−424.

Lynn, R., & Mikk, J. (in press). National differences in intelligence and educational attainment. *Intelligence*.

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

Lynn, R., & Vanhanen, T. (2006). *IQ and global inequality.* Augusta, GA: Washington Summit Books.

Marushiakova, E., & Popov, V. (2001). Historical and ethnological background. In W. Guy (Ed.), *Between past and future: the Roma of Central and Eastern Europe,* (pp. 33−53). Hatfield, England: University of Hertfordshire Press.

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

Oberauer, K., Suss, H. -M., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, *31*, 167−193.

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

Pearson, R. (1985). *Anthropological glossary.* Malabar, FL: Kreiger.

Raven, J., et al. (1990). *Raven manual research supplement 3: American and international norms*, (2nd ed.). Oxford, England: Oxford Psychologists Press.

Raven, J. C., Court, J. H., & Raven, J. (1995). *Raven manual: Coloured Progressive Matrices.* Oxford, England: Oxford Psychologists Press.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual: Standard Progressive Matrices.* Oxford, England: Oxford Psychologists Press.

Rushton, J. P. (1992). Cranial capacity related to sex, rank, and race in a stratified random sample of 6325 U.S. military personnel. *Intelligence*, *16*, 401−413.

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

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*, 235−294.

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., & Bons, T. A. (2004). Construct validity of Raven's Advanced Progressive Matrices for African and Non-African engineering students in South Africa. *International Journal of Selection and Assessment*, *12*, 220−229.

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

Rushton, J. P., Skuy, M., & Fridjohn, P. (2003). Performance on Raven's Advanced Progressive Matrices by African engineering students. *Intelligence*, *31*, 123−137.

Save the Children (2001). *Denied a future? The right to education of Roma, Gypsy and Travelers children.* London: Save the Children Fund.

Sowell, T. (2004). *Affirmative action around the world: An empirical study.* New Haven: Yale University Press.

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

Sternberg, R. J., Grigorenko, E. L., Ngrosho, D., Tantufuye, E., Mbise, A., Nokes, C., et al. (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., et al. (2001). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence*, *29*, 401−418.

Templer, D. I., & Arikawa, H. (2006). Temperature, skin color, per capita income, and IQ: An international perspective. *Intelligence*, *34*, 121−139.

te Nijenhuis, J., Tolboom, E. A., Resing, W. C., & Bleichrodt, N. (2004). Does cultural background influence the intellectual performance of children of immigrant groups? *European Journal of Psychological Assessment*, *20*, 10−26.

Whetzel, D. L., & McDaniel, M. A. (2006). Prediction of national wealth. *Intelligence*, *34*, 449−458.