

# Survey of Expert Opinion on Intelligence: Causes of International Differences in Cognitive Ability Tests

Heiner Rindermann<sup>1\*</sup>, David Becker<sup>1</sup>, Thomas Coyle<sup>2</sup>

<sup>1</sup>Chemnitz University of Technology, Germany, <sup>2</sup>Department of Psychology, University of Texas at San Antonio, USA

*Submitted to Journal:*  
Frontiers in Psychology

*Specialty Section:*  
Educational Psychology

*ISSN:*  
1664-1078

*Article type:*  
Original Research Article

*Received on:*  
18 Dec 2015

*Accepted on:*  
05 Mar 2016

*Provisional PDF published on:*  
05 Mar 2016

*Frontiers website link:*  
[www.frontiersin.org](http://www.frontiersin.org)

*Citation:*  
Rindermann H, Becker D and Coyle T (2016) Survey of Expert Opinion on Intelligence: Causes of International Differences in Cognitive Ability Tests. *Front. Psychol.* 7:399.  
doi:10.3389/fpsyg.2016.00399

*Copyright statement:*  
© 2016 Rindermann, Becker and Coyle. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

This Provisional PDF corresponds to the article as it appeared upon acceptance, after peer-review. Fully formatted PDF and full text (HTML) versions will be made available soon.

Header: Experts on national IQs

February 22, 2016

## **Survey of Expert Opinion on Intelligence: Causes of International Differences in Cognitive Ability Tests**

Heiner Rindermann<sup>1</sup>, David Becker<sup>1</sup> & Thomas R. Coyle<sup>2</sup>

<sup>1</sup>Department of Psychology, Chemnitz University of Technology, Germany

<sup>2</sup>Department of Psychology, University of Texas at San Antonio, United States

for *Frontiers in Psychology*

Word count: 5067 words (main text), 52 references, 177 words abstract, total 7339

### **Addresses**

*Prof. Dr. Heiner Rindermann* (corresponding author)  
Department of Psychology, Chemnitz University of Technology  
Wilhelm-Raabe-Str. 43, D-09107 Chemnitz, Germany  
Tel.: +49 371 531 38715, Fax +49 371 531 838715  
E-mail: heiner.rindermann@psychologie.tu-chemnitz.de

*David Becker*  
Department of Psychology, Chemnitz University of Technology  
Wilhelm-Raabe-Str. 43, D-09107 Chemnitz, Germany  
E-mail: david.becker@s2009.tu-chemnitz.de

*Prof. Dr. Thomas Coyle*  
University of Texas at San Antonio, Department of Psychology  
One UTSA Circle, San Antonio, TX 78249, USA  
Tel.: +1 210 458 7407; Fax: +1 210 458 5728  
Email: thomas.coyle@utsa.edu

### **Author contributions**

All three authors developed the study concept. David Becker programmed the web-based questionnaire and carried out the survey. David Becker and Heiner Rindermann performed the data analysis. David Becker drafted an initial version of the manuscript. Heiner Rindermann and Thomas Coyle provided revisions. All authors approved the final version of the manuscript for submission.

**Abstract**

Following Snyderman and Rothman (1987, 1988), we surveyed expert opinions on the current state of intelligence research. This report examines expert opinions on causes of international differences in student assessment and psychometric IQ test results. Experts were surveyed about the importance of culture, genes, education (quantity and quality), wealth, health, geography, climate, politics, modernization, sampling error, test knowledge, discrimination, test bias, and migration. The importance of these factors was evaluated for diverse countries, regions, and groups including Finland, East Asia, Sub-Saharan Africa, Southern Europe, the Arabian-Muslim world, Latin America, Israel, Jews in the West, Roma (gypsies), and Muslim immigrants. Education was rated by  $N=71$  experts as the most important cause of international ability differences. Genes were rated as the second most relevant factor but also had the highest variability in ratings. Culture, health, wealth, modernization, and politics were the next most important factors, whereas other factors such as geography, climate, test bias, and sampling error were less important. The paper concludes with a discussion of limitations of the survey (e.g., response rates and validity of expert opinions).

Σ 177 words

**Keywords:** Intelligence; Survey; Experts; cross-national differences.

# Survey of Expert Opinion on Intelligence: Causes of International Differences in Cognitive Ability Tests

## 1 Introduction

Cognitive ability is understood as the ability to think (intelligence), the disposition of knowledge (the store of true and relevant knowledge), and the intelligent use of this knowledge. The current study focuses on international differences in cognitive ability, based on test scores from different countries. International differences in cognitive ability have been estimated using student assessment studies (e.g., TIMSS, PISA, PIRLS; OECD, 2013) and psychometric IQs (Lynn & Vanhanen, 2012).<sup>1</sup> These international differences in cognitive ability are strongly related to country-level differences in wealth, health, democracy, lifespan, education, innovation, and gross-national product (e.g., Hanushek & Woessmann, 2015; Rindermann, Kodila-Tedika, & Christainsen, 2015).

Differences between countries with the lowest and highest ability levels are large. For example, in TIMSS 2011, 4<sup>th</sup> grade Yemeni pupils achieved 209 student assessment study (SAS) points, whereas South Korean pupils achieved 587 SAS points. If SAS points are converted to IQ points, the Yemeni would have an IQ of 56 and the Koreans would have an IQ of 113, a difference of 11 years of schooling.<sup>2</sup> Psychometric IQ studies show similar results. For example, Malawi has an estimated IQ of 60, whereas Singapore and Hong Kong have estimated IQs around 108, a difference that translates into SAS≈233 and 555 or 16 years of schooling.<sup>3</sup>

---

<sup>1</sup> PISA (Programme for International Student Assessment, internationally, since 2000), PIRLS (Progress in International Reading Literacy Study, since 2001), and TIMSS (Trends in International Mathematics and Science Study, since 1995).

<sup>2</sup> The conversion transforms the SAS-scale ( $M=500$ ,  $SD=100$ , in developed countries) to an IQ scale ( $M=100$ ,  $SD=15$ ). We assumed an increase of 35 SAS points, or 3 IQ points per year of schooling (Rindermann, 2011; Winship & Korenman, 1999).

<sup>3</sup> The comparability of student assessments and intelligence tests was analyzed in Rindermann and Baumeister (2015) and Rindermann (2007).

Cross-country differences in cognitive ability form geographically contiguous areas with smooth transitions. The lowest ability levels are found in Sub-Saharan Africa, the highest levels in East Asia, and moderate to high levels in Europe and other developed countries. The ability levels span a range of 3 *SDs* in IQ. Such a large range has raised questions about the validity, measurement, and causes of the differences (e.g., Hopmann, Brinek, & Retzl, 2007; Wicherts, Dolan, & Maas, 2010). Discussed causes have included education, modernization, politics, wealth, culture, and evolutionary-genetic factors (e.g., Lynn & Vanhanen, 2012).

Different factors have been stressed for different regions, nations, and cultures. These factors include the quality of educational and support systems (e.g., school health care, families) in Finland (Döbert, Klieme, & Sroka, 2004), low educational levels in Arabian countries (UNDP, 2003), a culture of effort and achievement in East-Asia (Helmke & Hesse, 2002), health issues in sub-Saharan Africa (Glewwe & Kremer, 2006), and evolutionary-genetic factors in sub-Saharan Africa and East-Asia (e.g., Kanazawa, 2004; Lynn, 1990; Rushton, 2004). Genetic factors have provoked particularly heated disputes, but even the mere description of international ability differences has been contentious (e.g., Coyle, Johnson & Rindermann, 2013; Hunt, 2012; Sternberg, 2013).

One solution to reduce conflict would be to survey experts – scientists who have conducted research on cognitive ability and who have an informed opinion about group differences. Such a solution was adopted in the current study, which examined expert opinions on international ability differences. It is possible that expert opinions may closely approximate the truth. However, majority itself is no sure indicator for truth, especially in a field with zeitgeist or political pressures that favor a particular viewpoint or opinion (e.g., Jussim, Crawford, Anglin, & Stevens, 2015).

But an expert survey shows two advantages: First, according to the Spearman-Brown prediction formula (1910), increases in the items being analyzed (e.g., test items or expert ratings) will increase the reliability of the final averaged result. The truth value of expert opinions was first described by Francis Galton (1907), who asked stockbreeders

---

On the comparability of student assessment and intelligence test scales see Rindermann & Baumeister (2015) and Rindermann (2007).

and butchers at a cattle market to estimate the weight of an ox. Galton analyzed the estimates and found that the mean value of all results was correct within 1 per cent of the real value. That is, the average opinion of experts was very close to the real answer. The accuracy of expert opinions may extend to the current study, which surveyed experts on international ability differences.

Second, in the current study, data collection procedures were designed to ensure anonymity. The anonymity was designed to reduce pressure for socially desirable responses, and to increase the likelihood of obtaining honest opinions. Opinions made in anonymity (without fear of retribution) may differ from public appraisals such as those reported in Gottfredson's (1994) "Mainstream Science on Intelligence", which was signed by more than 50 researchers.

Thirty years ago, Snyderman and Rothman (1987, 1988) surveyed experts on the concept, measurement, and heritability of intelligence, and also on race, class, and cultural differences in intelligence. In the current study, we surveyed experts on the causes of international ability differences. Such differences were not addressed by Snyderman and Rothman, but have received much recent attention (e.g. Hunt, 2012; Lynn & Vanhanen, 2012; Wicherts, Dolan, & Maas, 2010).

## **2 Method**

### *2.1 Survey: Expert Questionnaire on Cognitive Ability*

To examine expert opinions, we created a new questionnaire ("Expert Questionnaire on Cognitive Ability"). The questionnaire was partly based on the one developed by Snyderman and Rothman (1987, 1988), but addressed current research topics such as the Flynn effect and international ability differences. Additional questions addressed the definitions of terms and concepts (e.g., intelligence and cognitive ability); validity of tests (e.g., WAIS, SAT, PISA); estimation of genetic and environmental factors; intelligence and the media; genetic testing (e.g., DNA screening); and development of intelligence in global regions and populations. There were 62 main questions with follow-up questions and space for comments. The current study focused on questions concerning the causes of international ability differences.

## 2.2 Questions

The first general question was, “On causes of cross-national differences, what do you think are the reasons for large differences in cognitive ability and intelligence studies (including PISA, TIMSS etc.)?” This and subsequent questions were followed by the statement, “To address this question, please rate the factors below in importance (it is difficult to clearly distinguish them because they are somewhat connected, so please give only a rough estimate).” There were 15 factors, each rated on a scale from 0 to 100 in 5% steps, with “no answer” also possible. 0% would suggest that the factor explains 0% of international differences and is unimportant; 100% would suggest that the factor explains all of the differences and is highly important. The 15 factors were culture (e.g., religion, traditions), genes, educational quantity, educational quality, wealth, health, geography, climate, politics, modernization, sampling error, test knowledge, discrimination, test bias, and migration.

To obtain additional information, the same questions were asked for the following scientifically or politically important countries and groups:

- Finland: “One of the countries with the highest scores on international assessments (especially PISA) is Finland. Why do you think this is the case?”
- East Asia: “The region with the highest scores on international assessments is East-Asia (Japan, Korea, the Chinas). Why do you think this is the case?”
- Sub-Saharan Africa: “The region with the lowest scores on international assessments is Sub-Saharan Africa (Nigeria, Kenya or South Africa). Why do you think this is the case?”
- Southern Europe (compared to North-Western-Middle Europe): “Results in Southern Europe (Portugal, Italy or Greece) are lower than results in North-Western-Middle Europe (United Kingdom, France, Netherlands, Finland or Germany). Why?”
- Arabian-Muslim world (compared to North-Western-Middle Europe): “Results in the Arabian-Muslim world (Egypt, Saudi Arabia or Iran) are lower than results in North-Western-Middle Europe (United Kingdom, France, Netherlands, Finland or Germany). Why?”
- Latin America (compared to North America): “Results in Latin America (Mexico, Brazil or Argentina) are lower than results in USA or Canada. Why?”

- Israel (compared to North-Western-Middle Europe): “Results in Israel are lower than results in North-Western-Middle Europe. Why?”
- Jews in the Western world: “Results of Jews in the Western world are at the top. Why?”
- Roma: “Results of Roma (gypsies) in Europe are comparatively low. Why?”
- Muslim immigrants: “Results of immigrants from the Middle East (Arabian and Muslim countries) in the Western world are not at the top. Why?”

Each country or group was rated on the same 15 factors used to evaluate general cross-national differences, using the same scales (0 to 100), questions, and instructions. In addition, comments could be given for each factor being rated.

### 2.3 Experts

Notice of the study was emailed to experts who published articles on or after 2010 in journals on intelligence, cognitive abilities, and student achievement. The journals included *Intelligence*, *Cognitive Psychology*, *Contemporary Educational Psychology*, *New Ideas in Psychology*, and *Learning and Individual Differences*. Notice of the study was also emailed to members of the International Society for Intelligence Research (ISIR), and posted to the web site for the International Society for the Study of Individual Differences (ISSID). ISIR and ISSID support intelligence research and host professional conferences with intelligence researchers. Finally, notice of the study was announced at the 2013 ISIR conference in Melbourne, Australia. A total of 1,345 people received an email invitation. An expert was defined as a person who had published on cognitive ability or who had attended intelligence conferences and presented research. Compared to Snyderman and Rothman (1988, pp. 46-49), our selection criteria were based more on publications in specific scientific journals and less on membership in scientific organizations. In addition, we used email and a web-based survey rather than traditional mail and paper-pencil surveys.

We received a total of 265 responses from May 2013 to March 2014, at which time the survey was closed. The response rate was 20% of all invitations. The present article focused on cross-national differences and concerned questions towards the end of the survey. These questions were answered by 71 respondents. By comparison, there were 1,020 invitations and 661 participants (65%) in the Snyderman and Rothman study.

Respondents were 83% male and 17% female, and had a mean age of 49 years ( $SD=14.87$ ,  $median=47$  years). Respondents were mainly from the United States (38%), Germany (17%), Scandinavia (9%), the UK and Ireland (8%), Spain (7%), Canada (6%), and Australia/New Zealand (6%). The rest were from Latin America, Austria, Switzerland, Italy, Russia, The Netherlands, Greece, Portugal, and China. Around 20% had at least some self-rated Jewish heritage. The religious background of the respondents' families was about one third Catholic, one third Protestant, and one third with no affiliation. About 6% of respondents were Jewish (religion), and 62% professed no religious affiliation.

Participants worked in the fields of psychology (80%), education (8%), biology (5%), economics (3%), sociology (2%), and physical anthropology (2%). 87% had a Ph.D. All were scientists (i.e., no journalists). 67% were tenured faculty, 24% non-tenured faculty, and 6% students. Other participants (about 3%) were not categorized or worked in non-academic research institutes.

The majority of respondents did research on the “nature of intelligence/cognitive abilities”. The percentage of respondents that studied group differences in ability and related topics was 55% for “group differences”; 33% for “consequences for economics, society, and culture”; 26% for “bias in ability tests”; 25% for “cross-cultural comparisons”; and 25% for “FLynn effect”. Respondents published an average of 47 papers ( $median=20$  papers), and reported high  $h$ -indexes (mean Scopus  $h=17$ ,  $median h=11$ ; mean Harzing  $h=22$ ,  $median h=17$ ). The published papers had passed peer review, a requirement for inclusion in Scopus  $h$ -index counts. The counts substantiate the high quantity and quality of research published by respondents.

One limitation of the study can be seen in the small sample and low response rates. The sample consisted of 71 respondents, which is small compared to Snyderman and Rothman's sample of 661 respondents (of 1,020 invitations). In addition, self-selection of experts could have biased the results.

We attempted to increase response rates by using an Internet survey, emailing invitations (and reminders), and announcing the survey at intelligence conferences. Despite these measures, response rates were still low. The low response rates may be attributed to the length of the survey (which took about 40 to 90 minutes to complete), self-censorship, or fear of addressing a controversial subject (despite assurances of

anonymity). The low response rates may also reflect a paucity of experts on intelligence and international differences in cognitive ability. There may be 20 to 50 scientists who study international differences in intelligence. Based on this estimate, the number of respondents (71 people) may exceed the number of scientists who study the topic! Because the aim of the survey was to obtain expert opinions on the research questions, our view is that participation of people with only tangential knowledge of the subject matter could distort answers more than low response rates attributable to self-selection by experts.

One researcher suggested in an email that only politically biased researchers would respond to the questionnaire, especially given its length. In contrast to the speculation, three researchers sometimes labeled as “conservative” in Internet articles and Wikipedia refused to participate due to lack of expertise, survey length, or disapproval of opinion surveys as a way of finding truth. Based on several comments of researchers with diverse political background we have no hint of biased participation or refusal of participation.

#### 2.4 *Procedure*

The Internet based survey was created with the free software, LimeSurvey. Invitees received an invitation email with instructions on how to obtain a participation code, which could not be linked to the respondents' identities. Participants were told, “All questionnaires will be anonymously treated. No individual results will be published or given to other persons or the media. Results will be presented in an anonymous manner.” Only people who received a participation code could participate. After two weeks, invitees who had not responded received a reminder email.

#### 2.5 *Statistical analysis*

Experts could assign a percentage to all 15 factors (causes) for each question. Logically, the percentages across all factors should sum to 100%. However, experts did not always rate all factors, and their ratings did not always sum to 100%. To address variation in percentages across participants (and questions), we analyzed only responses from experts who rated at least 5 of 15 factors or whose ratings summed to at least 75%. The sum of the ratings of 5 factors usually explained more than the half of the international differences. In addition, we replaced missing ratings with 0%, and standardized the sum

of the ratings (within each question) to 100%. Replacing missing values with a score of 0% was based on the assumption that experts who did not rate a factor judged that factor as having no influence on ability differences.

The adjustments corrected for biases across raters and questions. For example, the average sum of items (i.e., factors) for the first global question (i.e., causes of cross-national differences) was 440%. With the corrections, the average sum of items for the first global question (and all other questions) was 100%. Without the corrections, the average sum of items for each question would deviate from 100%, and would prevent comparisons across questions.

### 3 Results

#### 3.1 Causes of general cross-national differences

*Table 1 around here please*

Seventy-one experts rated possible causes of cross-national differences in cognitive ability based on psychometric IQs and student assessment studies (e.g., PISA, PIRLS, TIMSS). Genes were rated as the most important cause (17%), followed by educational quality (11.44%), health (10.88%), and educational quantity (10.20%) (Table 1). The sum of both education factors yielded the highest rating (21.64%). Of all factors, genes had by far the largest standard deviation ( $SD=23.85$ ; all other factors,  $SD<10$ ), indicating disagreement about the importance of genetic influences. Only 5 of 71 experts (7%) who responded to the genetic item thought that genes had no influence. If non-responses to the genetic item are converted to 0% (4 additional experts), 13% of experts doubted any genetic influence. The frequency of zero-percentage-ratings was larger for genes than for culture or education (about 1%), but experts who believed that genes had no influence were a minority: Around 90% of experts believed that genes had at least some influence on cross-national differences in cognitive ability.

Items with lower percentages ( $<10\%$ ) included wealth, culture, and modernization (7% to 9%). Methodological bias factors (sampling error, test knowledge, test bias) were rated as less important (3% to 6%, together 11.78%).

### 3.2 *Causes for the ability levels of single countries and groups*

The *Finnish miracle* refers to the top student test results in Finland (Simola, 2005). The top results in Finland have been attributed to progressive educational methods.

However, traditional factors and declines in them (e.g., discipline and teacher authority) have been associated with recent drops in test results (Sahlgren, 2015). In the current study, experts attributed the top results in Finland to high educational quality (19.11%), followed by genetic-evolutionary factors (14.89%) and educational quantity (12.71%). The two educational factors combined totaled 31.87%. Educational quality was rated higher in Finland than in any other country.

*East Asian countries* attain high student test results in nearly all student assessment and psychometric IQ studies. Early test results from East Asia were attributed to better and more efficient instruction (Stigler et al., 1999), culture (Confucian achievement orientation; Helmke & Hesse, 2002), and evolutionary-genetic factors (Kanazawa, 2004; Lynn, 1990; Rushton, 2004). In the current study, experts attributed the high results in East Asia to genes (19.89%), education (quantity: 16.87%, quality: 14.72%), and culture (14.49%). Compared to other countries and regions, East Asia had the strongest ratings for educational quantity and test knowledge.

*Sub-Saharan Africa* has the lowest student test results (in participating countries), a finding confirmed by psychometric IQs and Piagetian tasks (Hallpike, 1978; Lemos, 1974; Rindermann, 2013; Rindermann, Falkenhayn & Baumeister, 2014). Similar to discussions of low intelligence in the US, the findings for Sub-Saharan Africa are highly contentious, especially when factors other than the environment are examined (e.g. Nyborg, 2003; Segerstråle, 2000). Factors that have been implicated in the low results include health, wealth, evolution, political problems (corruption affecting educational means), modernization, and education (e.g. Glewwe & Kremer, 2006; Lynn, 1990). Measurement issues were also mentioned (e.g. Wicherts et al., 2010; Wober, 1969). In the current study, experts attributed the low results in Sub-Saharan Africa to genetic-evolutionary factors (18.58%), followed by educational quality (12.27%), health (11.73%), and educational quantity (11.60%). The two educational factors together had the strongest rating (23.87%), and health had a high rating compared to other regions and countries (12%). While genes were rated as the most important single factor, there was considerable diversity of opinion: 10 of 60 experts gave genes a rating of zero

(17%), and the standard deviation in ratings for genes was the highest of all factors ( $SD=24.88$ ; all other factors:  $SD<10$ ). Similar to other countries and regions, Sub-Saharan Africa had low ratings for discrimination and methodological problems (sampling error, test bias, test knowledge).

*Southern Europe* attains low student test results compared to Central, West, and Northern Europe, and to North America and Australia/New Zealand.<sup>4</sup> The low results have been implicated in recent economic events (e.g., European debt crisis; Vanhanen, 2013). In the current study, experts attributed the low results in Southern Europe to genes (16.78%), educational quality (15.07%), and culture (12.28%). Compared to other countries and regions, Southern Europe had the highest rating for political factors (9.38%).

*Arab-Muslim countries* have low student test results and psychometric IQs, and also low levels of educational and scientific achievements, compared to other regions and countries (e.g., Arab Human Development Report; UNDP, 2003). Even Arab students in engineering (which draws high ability students) typically achieve only average test results (Rindermann, Baumeister, & Gröper, 2014). In the current study, experts attributed the low results in Arab-Muslim countries to genetic-evolutionary factors (17.09%), followed by culture (14.70%), educational quality (14.41%), and educational quantity (13.00%). Compared to other regions, low modernization received the highest rating for Arab-Muslim countries (6.58%), a finding consistent with the Arab Human Development Report (UNDP, 2003).

*Latin America* lags behind other regions in student assessment results, a finding that has been attributed to culture (e.g., Harrison, 2006). In the current study, experts attributed the low results in Latin America to genetic (16.42%) and educational factors (quality: 15.44%, quantity: 13.46%). The two educational factors combined had the strongest rating (28.90%). Culture received a moderate rating (11.34%), a finding similar to other countries (Table 1).

*Israel* is a puzzling case due to its relatively low test results compared to Jews in the West. The question was: “Results in Israel are lower than results in North-Western-

---

<sup>4</sup> However, the picture is not uniform. The northern regions of Italy and Spain have results similar to those of other Western countries (Lynn, 2010, 2012; OECD, 2007).

Middle Europe. Why?” Lynn (2011) attributed the low results in Israel to diverse ethnic groups (e.g., Ashkenazim, Sephardim, Mizrahim, Ethiopian Jews, Palestinian Arabs), some of which lower the overall ability level. In the current study, experts attributed the Israeli results to the usual factors (genes: 20.49%; educational quality: 10.69%; educational quantity: 8.94%), with genes being rated slightly more important than both educational factors combined (19.63%). Compared to other countries and regions, education had the lowest rating, whereas immigration had the highest rating (8.20%).

*Jews in the West* show high student test results and also exceptional intellectual, scientific, and cultural achievements (e.g., Haag, 1969; Lipset & Raab, 1995). In the current study, experts attributed the high test results of Jews to genes (28.36%), which were rated more strongly than the two educational factors combined (quantity: 10.79%, quality: 13.38%; combined: 24.17%). Genetic factors had the highest rating compared to all other groups and countries. However, as in the cross-country analyses, there was disagreement about the importance of genes, with variability in ratings being higher for genes than for other factors ( $SD = 33.10$ ; all other factors:  $SD < 16$ ).

*Roma* (Romani people, gypsies with different subgroups) are the lowest scoring long-term natives in Europe (e.g., Cvorovic, 2014). The low test results have been attributed to low levels of education, health, wealth, and modernization, as well as discrimination and genetic factors (Cvorovic, 2014; Rushton, Cvorovic, & Bons, 2007). The expert survey revealed no special pattern of results. Experts attributed the low results of the Roma primarily to genetic factors (21.35%), followed by culture (15.16%) and education (quality: 14.73%, quantity: 14.15%). The two educational factors combined were seen as most influential (28.88%).

*Immigrants from the Middle East* (Arabian and Muslim countries) generally show low student test results and psychometric IQs, which are a third to one standard deviation lower than other groups (e.g. Dronkers, Velden, & Dunne, 2012; te Nijenhuis & Flier, 2001). The low test results of Middle Eastern immigrants have been attributed to culture, education, language, and discrimination (Dronkers et al., 2012). In the current study, experts attributed the low test results primarily to genetic factors (23.01%) and culture (17.26%). Discrimination was rated as relatively unimportant for Middle Eastern immigrants (2.27%) and for other groups and countries (1.25%). The

two educational factors combined were rated as most important (quality: 13.30%, quantity: 12.14%, combined: 25.44%).

### *3.3 Causes for the ability levels of single countries and groups – average ratings*

The strongest rated factor across all countries, regions, and groups was genes-evolution (19.72%), followed by education quality (14.69%), culture (13.71%), and educational quantity (13.60%; Table 1, last row). These four factors were followed by health (7.32%), wealth (7.27%), politics (5.56%), and modernization (4.90%). Other factors received weak ratings (e.g., geography, current climate, migration, discrimination), including methodological factors (e.g., sampling error, test knowledge, test bias). Nature and nurture compared, genes were rated as the most important factor, but the two educational factors combined (quality and quantity), and all environmental factors together (e.g., education, culture, wealth, health), were rated more strongly than genes.

Cultural factors were rated as more important for single countries than for general cross-country comparisons (13.71% vs. 7.35%). In addition, compared to cross-country comparisons, single countries showed higher ratings for genes (19.72% vs. 16.99%), educational quantity (13.60% vs. 10.20%), and educational quality (14.69% vs. 11.44%). In contrast, single countries showed lower ratings for wealth, health, modernization, and test bias (total of sampling error, test knowledge, and test bias: 6.26% single-country vs. 11.78% cross-country), suggesting that general cross-country comparisons reflect global impressions that may not hold for single countries.

### *3.4 Comments*

There were few entries in the comment sections. Two experts noted that the results for Israel depended on a mix of ethnicities and climate, without detailed explanation. Other experts noted that strong discrimination may be a cause of Jewish high intelligence, and that the causes of international ability differences are interconnected. Another expert stressed that Finns have the “fastest simple visual reaction times of any contemporary Western population”, and referred to a publication.

## 4 Discussion

Experts in cognitive ability rated possible causes of international differences in student assessment and psychometric IQ test results. Ratings were obtained for cross-national differences and single countries using a percentage scale ranging from 0% to 100% (higher percentages = more important). The expert survey revealed important results: Methodological factors (sampling error, test bias, test knowledge) were weakly rated (cross-national: 11.78%; single countries' average: 6.26%), as was discrimination (cross-national: 2.10%, single countries: 1.23%). In contrast, educational factors (quality and quantity) and genes were strongly rated. The low ratings for methodological factors suggest that international assessments were perceived to be valid indicators of cognitive ability and cross-country patterns.

Experts rated the two *educational factors* together (quantity and quality) as the most important cause of international differences in cognitive ability (cross-national: 21.64%, single countries' average: 28.29%). Weaker ratings were given for environmental factors such as health (cross-national: 10.88%, single countries: 7.32%), wealth (cross-national: 8.96%, single countries: 7.28%), modernization (cross-national: 7.19%, single countries: 4.91%), and politics (cross-national: 4.77%, single countries: 5.56%). The sum of all *environmental factors* explained more than half of the international ability differences (cross-national: 53.44%; single- countries' average: 53.36%).

The relative importance of environmental factors does not mean that genetic factors were seen as irrelevant. Based on expert opinions, the *genetic-evolutionary factor* was the single most important cause of international differences in cognitive ability (cross-national: 16.99%, single-country: 19.72%): Experts attributed about one-sixth to one-fifth of international ability differences to genes. While the rated impact of genes was remarkable, it was still well below the rated impact of environmental factors (around 50%). In addition, disagreement among experts (based on *SDs* in ratings) was much higher for genes than for environmental factors. Other factors such as geographic and climate differences were rated as negligible influences on international ability differences. Immigration was rated as important only for Israel.

While both camps, the nature as well as the nurture camp, are supported by the results of the expert survey, the scientifically more interesting result is the support for genetic explanations. The environmental theory is in research not questioned, its

corroboration by an expert survey is not astonishing. More important is the support for the frequently hotly disputed genetic explanations. Assuming that the survey is representative of expert opinions, genetic factors should receive more attention in future research and public debates. To fairly consider different hypotheses, future research should incorporate procedures (e.g., rules for methods of argumentation) that reduce zeitgeist or political pressures that may bias responses on controversial issues (e.g., Jussim, Crawford, Anglin, & Stevens, 2015; Segerstråle, 2000).

This finally leads to some limitations of the presented study: The main limitation can be seen the low response rate (20%). Although low response rates can produce unreliable results, low response rates are less problematic in the current study, which was designed to solicit *informed* opinions about ability differences. Because experts who responded to the survey had conducted research in relevant areas (e.g., cognitive ability, group differences, test bias, cross-cultural comparisons, Flynn effects), they were in a good position to provide informed opinions of ability differences. Thus, although response rates were low, the responses were likely to reflect the opinions of people who had expertise on the topic being evaluated.

A related limitation is self-selection. Only a small number of experts responded to the invitation, and the responses of these experts may be unrepresentative and biased. Such problems are common to surveys, but might be less relevant in the current study, which was designed to attract experts in a narrow domain. (There are perhaps no more than 50 experts in the world who specialize in international ability differences.) The experts who completed the survey reported that they had expertise in intelligence and group differences. This suggests that the experts had self-selected into the survey based on relevant expertise, which should increase the validity of opinions. Based on our experience, we suggest that future surveys more carefully preselect who is invited to participate. In the current study, our aim was to recruit a broad sample of experts, including experts who might be reluctant to respond for various reasons, but not to exclude anybody who perceives him- or herself as an expert. In retrospect, our recruitment procedure was probably too broad, yielding many invitations but few responses, which lowered response rates.

A third limitation concerned the measurement of factors that cause international ability differences. Different factors were measured with different numbers of items.

Education was measured with two items, environmental factors with 11 items, and genetics with one item. To better estimate the importance of nature and nurture, a single binary question could be added to future surveys (e.g., “Which is more influential, genetic or environmental factors?”). In addition, the survey did not assess indirect influences. Additionally, genetic effects can be included in the rated environmental aspects health, wealth, institutional features and migration or in non-rated aspects of personality (Dawkins, 2008/1992; Spolaore & Wacziarg, 2013; Woodley, Rindermann, Bell, Stratford, & Piffer, 2014). Finally, culture influences via consanguineous marriages genetic factors (e.g. Tadmouri et al., 2009).

Further research is needed to identify the facets of the factors that contributed to expert opinions. The facets may include different aspects of educational quality (e.g., student discipline or class size) or different types of genes (e.g., generalist genes or genes for specific abilities).

It should be noted that an expert survey is both an *opinion instrument* and a possible *indicator of truth*. As an opinion instrument, the survey would be biased if the selection of experts were biased. In the current study, experts were selected if they had published on intelligence or attended an intelligence conference, increasing the likelihood that the experts had an informed opinion. The other aspect of an expert survey – that it is an indicator of truth – cannot be answered by this study. A single empirical study can contradict expert opinion, and the results of the current survey must be validated in future empirical research.

Finally, the expert survey did not address *the interdependence of factors*. Factors related to cognitive ability may influence each other in complex ways. For example, culture may influence education, genes, health, and cognitive ability, which in turn may influence wealth and economic development (e.g. Rindermann, Hoang & Baumeister, 2013; Rindermann et al., 2015). These interdependencies could be examined using path models and longitudinal datasets. Although expert surveys cannot replace empirical research, the current survey provides a snapshot of expert opinions, can give a more balanced view about the state of affairs in a research field and highlights targets for future studies (e.g., educational quality and genetic factors).

## References

- Coyle, Th., Johnson, W., & Rindermann, H. (2013, May 17). Open Letter – Letter of Protest on the comment of Robert Sternberg (2013) on Hunt (2012). *Perspectives on Psychological Science*, [http://pps.sagepub.com/content/8/2/187.abstract/reply#sppps\\_el\\_84](http://pps.sagepub.com/content/8/2/187.abstract/reply#sppps_el_84)
- Cvorovic, J. (2014). *The Roma: A Balkan underclass*. London: Ulster Institute for Social Research.
- Dawkins, R. (2008/1982). *The extended phenotype. The long reach of the gene*. Oxford: Oxford University Press.
- Döbert, H., Klieme, E., & Sroka, W. (Eds.). (2004). *Conditions of school performance in seven countries: A quest for understanding the international variation of PISA results*. Münster: Waxmann.
- Dronkers, J., Velden, R. v. d., & Dunne, A. (2012). Why are migrant students better off in certain types of educational systems or schools than in others? *European Educational Research Journal*, 11, 11-44.
- Galton, F. (1907). Vox populi. *Nature*, 75, 450-451.
- Glewwe, P., & Kremer, M. (2006). Schools, teachers, and education outcomes in developing countries. In E. A. Hanushek & F. Welch (Eds.), *Handbook of the economics of education* (II, pp. 945-1017). Amsterdam: North-Holland.
- Gottfredson, L. S. (1994). Mainstream science on intelligence. *The Wall Street Journal*, December, 13, A18.
- Haag, E. v. d. (1969). *The Jewish mystique*. New York: Dell.
- Hallpike, Ch. (1978). *Foundations of primitive thought*. Oxford: University Press.
- Hanushek, E. A., & Woessmann, L. (2015). *The knowledge capital of nations. Education and the economics of growth*. Cambridge, MA: MIT Press.
- Harrison, L. E. (2006). *The central liberal truth: How politics can change a culture and save it from itself*. New York: Oxford University Press.
- Helmke, A., & Hesse, H.-G. (2002). Kindheit und Jugend in Asien. [Childhood and youth in Asia.] In H.-H. Krüger & C. Grunert (Eds.), *Handbuch Kindheits- und Jugendforschung* [Handbook research on childhood and youth] (pp. 439-471). Opladen: Leske+Budrich.

- Hopmann, S. T., Brinek, G., & Retzl, M. (Eds.). (2007). *PISA zufolge PISA - PISA according to PISA*. Wien: LIT.
- Hunt, E. (2012). What makes nations intelligent? *Perspectives on Psychological Science*, 7, 284-306.
- Jussim, L., Crawford, J. T., Anglin, S. M. & Stevens, S. T. (2015). Ideological bias in social psychological research. In J. P. Forgas, K. Fiedler, & W. D. Crano (Eds.), *Social psychology and politics* (pp. 91-109). New York: Taylor & Francis.
- Kanazawa, S. (2004). General intelligence as a domain-specific adaptation. *Psychological Review*, 111, 512-523.
- Lemos, M. M. (1974). The development of spatial concepts in Zulu children. In John W. Berry & P. R. Dasen (Eds.), *Culture and cognition: Readings in cross-cultural psychology* (pp. 367-380). London: Methuen.
- Lipset, S. M., & Raab, E. (1995). *Jews and the New American Scene*. Cambridge: Harvard University Press.
- Lynn, R. (1990). Race differences in intelligence: A global perspective. *Mankind Quarterly*, 31, 255-296.
- Lynn, R. (2010). In Italy, north-south differences in IQ predict differences in income, education, infant mortality, stature, and literacy. *Intelligence*, 38, 93-100.
- Lynn, R. (2011). *The chosen people: A study of Jewish intelligence and achievement*. Whitefish: Washington Summit.
- Lynn, R. (2012). North-south differences in Spain in IQ, educational attainment, per capita income, literacy, life expectancy and employment. *Mankind Quarterly*, 52, 265-291.
- Lynn, R., & Vanhanen, T. (2012). *Intelligence. A unifying construct for the social sciences*. London: Ulster Institute for Social Research.
- Nyborg, H. (2003). The sociology of psychometric and bio-behavioral sciences. In H. Nyborg (Ed.), *The scientific study of general intelligence. Tribute to Arthur R. Jensen* (pp. 441-502). Oxford: Pergamon.
- OECD (2007). *PISA 2006. Volume 2: Data/Données*. Paris: OECD.
- OECD (2013). *PISA 2012 results: What students know and can do*. Paris: OECD.
- Rindermann, H. (2011). Increase of intelligence in childhood and youth. *Psychologie in Erziehung und Unterricht*, 58, 210-224.

- Rindermann, H. (2013). African cognitive ability: Research, results, divergences and recommendations. *Personality and Individual Differences*, 55, 229-233.
- Rindermann, H., & Baumeister, A. E. E. (2015). Validating the interpretations of PISA and TIMSS tasks: A rating study. *International Journal of Testing*, 15, 1-22.
- Rindermann, H., Baumeister, A. E. E., & Gröper, A. (2014). Cognitive abilities of Emirati and German engineering university students. *Journal of Biosocial Science*, 46, 199-213.
- Rindermann, H., Hoang, Q. S. N. & Baumeister, A. E. E. (2013). Cognitive ability, parenting and instruction in Vietnam and Germany. *Intelligence*, 41, 366-377.
- Rindermann, H., Kodila-Tedika, O., & Christainsen, G. (2015). Cognitive capital, good governance, and the wealth of nations. *Intelligence*, 51, 98-108.
- Rindermann, H., Falkenhayn, L., & Baumeister, A. E. E. (2014). Cognitive ability and epistemic rationality: A study in Nigeria and Germany. *Intelligence*, 47, 23-33.
- Rushton, J. Ph. (2004). Placing intelligence into an evolutionary framework or how g fits into the r-K matrix of life history traits including longevity. *Intelligence*, 32, 321-328.
- Rushton, J. Ph., Cvorovic, J., & Bons, T. A. (2007). General mental ability in South Asians: Data from three Roma (Gypsy) communities in Serbia. *Intelligence*, 35, 1-12.
- Sahlgren, G. H. (2015). *Real Finnish lessons. The true story of an education superpower*. London: Centre for Policy Studies.
- Segerstråle, U. (2000). *Defenders of the truth. The sociobiology debate*. New York: Oxford University Press.
- Simola, H. (2005). The Finnish miracle of PISA: Historical and sociological remarks on teaching and teacher education. *Comparative Education*, 41, 455-470.
- Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist*, 42, 137-144.
- Snyderman, M., & Rothman, S. (1988). *The IQ-Controversy, the Media and Public Policy*. New, Brunswick, New Jersey: Transaction.
- Spearman, Ch. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, 3, 271-295.

- Spolaore, E., & Wacziarg, R. (2013). How deep are the roots of economic development? *Journal of Economic Literature*, 51, 325-369.
- Sternberg, R. J. (2013). The intelligence of nations – smart but not wise: Comment on Hunt (2012). *Perspectives on Psychological Science*, 8, 187-189.
- Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study*. Washington: National Center for Education Statistics.
- Tadmouri, G. O., Nair, P., Obeid, T., Ali, M. T. A., Khaja, N. A., & Hamamy, H. A. (2009). Consanguinity and reproductive health among Arabs. *Reproductive Health*, 6, 1-9.
- te Nijenhuis, J., & Flier, H. (2001). Group differences in mean intelligence for the Dutch and Third World immigrants. *Journal of Biosocial Science*, 33, 469-475.
- UNDP (United Nations Development Programme) & Arab Fund for Economic and Social Development (2003). *Arab Human Development Report 2003. Building a knowledge society*. New York: UNDP.
- Vanhanen, T. (2013). Differences in national IQs behind the eurozone debt crisis? *Journal of Social, Political and Economic Studies*, 38, 3-15.
- Wicherts, J. M., Dolan, C. V., & Maas, H. L. J. (2010). A systematic literature review of the average IQ of sub-Saharan Africans. *Intelligence*, 38, 1-20.
- Wober, M. (1969). Distinguishing centri-cultural from cross-cultural tests and research. *Perceptual and Motor Skills*, 28, 488.
- Woodley, M. A., Rindermann, H., Bell, E., Stratford, J., & Piffer, D. (2014). The relationship between Microcephalin, ASPM and intelligence: A reconsideration. *Intelligence*, 44, 51-63.



Table 1: Means for each question and scale

	Culture	Genes	Education quantity	Education quality	Wealth	Health	Geography	Current climate	Politics	Modernization	Sampling error	Test knowledge	Discrimination	Test bias	Migration
Cross-national	7	<b>17</b>	10	<b>11</b>	9	<b>11</b>	3	2	5	7	4	6	2	3	3
Finland	10	<b>15</b>	<b>13</b>	<b>19</b>	9	10	2	2	6	6	2	3	1	1	3
East Asia	14	<b>20</b>	<b>17</b>	<b>15</b>	5	5	2	1	5	5	2	5	1	1	2
Sub-Sah. Africa	8	<b>19</b>	<b>12</b>	<b>12</b>	9	<b>12</b>	3	3	7	6	2	4	1	2	1
South Europe	<b>12</b>	<b>17</b>	11	<b>15</b>	8	6	3	3	9	6	2	4	1	1	3
Arab-Muslim	<b>15</b>	<b>17</b>	13	<b>14</b>	5	7	2	2	8	7	2	4	1	1	1
Latin America	11	<b>16</b>	<b>13</b>	<b>15</b>	8	8	2	2	7	6	2	4	1	1	1
Israel	<b>14</b>	<b>20</b>	9	<b>11</b>	6	7	3	2	8	4	2	3	1	1	8
Jews in West	<b>16</b>	<b>28</b>	11	<b>13</b>	8	6	1	1	3	4	3	3	2	1	2
Roma	<b>15</b>	<b>21</b>	14	<b>15</b>	7	7	1	1	3	6	1	3	2	1	2
Muslim immig.	<b>17</b>	<b>23</b>	12	<b>13</b>	6	5	1	1	4	5	2	3	2	3	2
Single average	<b>14</b>	<b>20</b>	<b>14</b>	<b>15</b>	7	7	2	2	6	5	2	3	1	1	2

Notes: Scale between 0 (no influence) and 100 (explains all). Values per row may not sum to 100% due to rounding error. For better

reading and comparison rounded to integer. Respondents per row, beginning with “Cross-national” and ending with “Muslim immigration”, are as follows: N=71, 62, 64, 60, 51, 52, 48, 46, 55, 49, and 46. The three most important factors per row are printed in bold. The bottom row (“Single average”) is the average rating across all single countries, regions, and groups (excluding cross-national general ratings in the first row).

## Appendix

The questions on international differences and their causes using the original question numbering.

- 40a On causes of cross-national differences, what do you think are the reasons for large differences in cognitive ability and intelligence studies (including PISA, TIMSS etc.)?
- 40b One of the countries with the highest scores on international assessments (especially PISA) is Finland. Why do you think this is the case?
- 40c The region with the highest scores on international assessments is East-Asia (Japan, Korea, the Chinas). Why do you think this is the case?
- 40d The region with the lowest scores on international assessments is sub-Saharan Africa. Why do you think this is the case?
- 40e Results in Southern Europe are lower than results in North-Western-Middle Europe. Why?
- 40f Results in the Arabian-Muslim world are lower than results in North-Western-Middle Europe. Why?
- 40g Results in South America are lower than results in North-America. Why?
- 40h Results in Israel are lower than results in North-Western-Middle Europe. Why?
- 40i Results of Jews in the Western world are at the top. Why?
- 40j Results of Roma (gypsies) in Europe are comparatively low. Why?
- 40k Results of immigrants from the Middle East (Arabian and Muslim countries) in the Western world are not at the top. Why?

### **For every single question the instruction was:**

“To address this question, please rate the factors below in importance (it is difficult to clearly distinguish them because they are somewhat connected, so please give only a rough estimate).”

### **To be rated casual factors were (in this order):**

Culture (religion, tradition, etc.); Genes (evolution); Education (quantity); Education (quality); Wealth; Health; Geography; Current climate; Politics; Modernization; Sampling error etc.; Test knowledge; Discrimination (-/+); Test bias (-/+); Migration (-/+)

**The rating scale was:**

01-05%; 06-10%; 11-15%; 16-20%; 21-25%; 26-30%; 31-35%; 36-40%; 41-45%; 46-50%; 51-55%; 56-60%; 61-65%; 66-70%; 71-75%; 76-80%; 81-85%; 86-90%; 91-95%; 96-100%; No answer.

Provisional