National IQs calculated and validated for 108 nations

Richard Lynn [a,*], Gerhard Meisenberg [b]

[a] University of Ulster, Coleraine, Northern Ireland, BT52 1SA, UK
[b] Ross University, Dominica

ARTICLE INFO

ABSTRACT

We estimate the validity of the national IQs presented by Lynn and Vanhanen (2002, 2006) by examining whether they are consistent with the educational attainment of school students in math, science and reading comprehension in 108 countries and provinces. The educational attainment scores in a number of studies are integrated to give EAs (educational attainments) for 86 countries and provinces that have measured IQs. The correlation of EA with measured IQ ($N = 86$ countries) is .917, and with measured + estimated IQ ($N = 108$ countries) is .907. Corrected for attenuation, $r = 1.0$. The quality of the data is evaluated, and the discrepancy between IQ and EA was greater for countries with low data quality, especially low quality of the EA data. There are no major anomalies in the relationship between IQ and EA in individual world regions. To some extent IQ and EA predict each other even within world regions. These results show that national IQs have a high degree of validity.

© 2010 Elsevier Inc. All rights reserved.

Contents

1. Introduction

National IQs for all countries in the world have been presented by Lynn and Vanhanen (2002, 2006). They have been welcomed by some as opening up a new field in the social

* Corresponding author.
   E-mail address: Lynnr540@aol.com (R. Lynn).

sciences in which IQ explains significant proportions of the variance in economic, demographic, and epidemiological outcomes at the country level, while a number of critics have dismissed national IQs as meaningless. Our objective in this paper is to test the validity of these national IQs by examining whether they are consistent with the educational attainment of school students in math, science and reading comprehension.

In their first study, Lynn and Vanhanen (2002, p.89) gave measured national IQs calculated from a variety of tests for 81 nations. These national IQs were measured from samples given a variety of intelligence tests and were calculated on a metric in which the British IQ is set at 100 (sd 15), and the IQs of other nations are calculated in relation to this standard. This has become known as the "Greenwich-IQ" metric, analogous to lines of longitude that are calibrated in relation to zero passing through Greenwich (a suburb of London). In addition, this study presented estimated IQs for a further 104 nations, giving a total of 185 nations and comprising all the nations in the world with populations over 50,000. The estimated IQs were obtained from the measured IQs of neighboring countries with culturally and racially similar populations. For example, there was no measured IQ for Pakistan, but there was a measured IQ (81) for India. It was assumed that the IQ in Pakistan would be approximately the same as in India, and hence Pakistan was assigned an IQ of 81.

The initial objective was to examine whether intelligence, measured as national IQ, contributes to national per capita income. This relationship was predicted based on repeated findings that intelligence is positively related to income among individuals. This latter relationship was shown for the United States by Jencks (1972), who calculated a correlation of .310 between IQ and income for men. This has been confirmed for Britain by Irwing and Lynn (2006), who reported data for a national sample whose intelligence was obtained at the age of 8 years and whose income was obtained at the age of 43 years. The correlations between IQ and income were .368 for men ($n = 1280$) and .317 for women ($n = 1085$). The hypothesis proposed by Lynn and Vanhanen (2002) was that nations can be considered as aggregates of individuals, and therefore that the positive relation between intelligence and income present among individuals should also hold for nations.

To test this hypothesis, Lynn and Vanhanen (2002) examined the relationship between measured national IQs of 81 nations and GDP (gross domestic product) in 1998. The correlation was .73. It was argued from this result that national IQs explained 53 per cent of the variance in national per capita income, and therefore that it provided a major contribution to a long-standing problem in development economics first raised by Adam Smith (1776) in his *Inquiry into the Nature and Causes of the Wealth of Nations*, and summarized by Landes (1999) as the problem of "Why some are so rich and some are so poor". The solution proposed by Lynn and Vanhanen (2002) to this question was that some are so rich partly because they have higher average IQs than those who are so poor. It was argued that this should be regarded as a causal relationship because it is an extension to nations of the established causal relationship of IQ to income among individuals.

This analysis was extended by Lynn and Vanhanen (2002) to all 185 nations in the world by using estimated IQs for nations for which no measured IQs were available. The correlation between national IQs for all 185 nations and per capita income measured as real GDP (Gross Domestic Product per capita, 1998) was .62. It was proposed that the lower correlation obtained with the 185 nations than with the 81 nations (.73) suggests some degree of inaccuracy in the estimated IQs.

This analysis was extended in a second study in which Lynn and Vanhanen (2006) gave measured IQs for 113 nations and estimated IQs for a further 79 nations, giving a total of 192 nations, comprising all the nations in the world with populations over 40,000. It was shown that for the 113 nations for which there were measured IQs, the correlation of national IQs with per capita income measured as GNI (Gross National Income at Purchasing Power Parity, 2002) was .68. For 189 nations the correlation of national IQs with per capita income measured as GDP (Gross Domestic Product per capita, 2000) was .64. Once again, the correlation of measured national IQs with per capita income (.68) is higher than the correlation using estimated national IQs (.64), suggesting some degree of inaccuracy in the estimated IQs.

These national IQs have evoked both interest and criticism. Some sceptics have dismissed them as "highly deficient" (Volken, 2003, p.411), "virtually meaningless" (Barnett & Williams, 2004, p.392), "technically inadequate… and meaningless" (Hunt & Sternberg 2006, pp. 133,136). For others, the calculation of national IQs has opened up a new field in which intelligence has explanatory power for a wide range of social phenomena. The calculation of national IQs and their correlates have been described by Rindermann and Ceci (2009, p. 551) as "a new development in the study of cognitive ability: following a century of conceptual and psychometric development in which individual and group (socioeconomic, age, and ethnic) differences were examined, researchers have turned their attention to national and international differences in cognitive competence. The goal is to use cognitive differences to understand and predict national differences in a variety of outcomes: societal development, rate of democratization, population health, productivity, gross domestic product (GDP), and wage inequality".

A number of investigators have reported a wide range of social, demographic and epidemiological correlates of national IQs. Lynn and Vanhanen's original claim that national IQs are correlated with per capita income at .73 has been refined by Meisenberg (2004) who has shown that the use of log GDP (1975–2003) increases the correlation based on 81 nations to .82. This correlation has been confirmed for 185 countries ($r = .65$) by Whetzell and McDaniel (2006), and for 152 countries ($r = .76$) by Morse (2008). Others have reported that Lynn and Vanhanen's national IQs are significantly correlated with rates of infant mortality (126 countries, $r = −.84$; Kanazawa, 2006); life expectancy (126 countries, $r = .75$ Kanazawa, 2006); total fertility rate (170 countries, $r = −.83$; Meisenberg, 2009); the incidence of HIV (165 countries, $r = −.48$, Rindermann & Meisenberg, 2009); the incidence of homicide (113 countries, $r = −.25$, Rushton & Templer, 2009); and skin color (129 countries, $r = .92$, Templer & Arikawa, 2006; see also Meisenberg, 2004, 2009).

These correlations can be regarded as validation of national IQs and suggest that national IQs have considerable

explanatory power. Nevertheless, to meet those critics who consider that "the concept of national IQ is meaningless" (Hunt and Sternberg (2006, pp. 133), we employ here one of the classical methods for establishing the validity of intelligence tests: examining whether they are correlated with educational attainment. As Matarazzo (1972, p.281) has noted: "thousands of studies have been published, in numerous languages throughout the world, attempting to demonstrate the validity of intelligence tests against academic performance in school".

2. National educational achievements (EAs) for 108 countries

We now present measures of educational achievement, defined by the scores of school students on international assessments of mathematics, science, and reading. These will be labeled "educational quotients" (EAs). Data are available for 108 nations. 86 of these nations also have measured IQs. This exercise extends and updates previous work by Rindermann (2007a, b).

2.1. General strategy

The major international school assessment studies are TIMSS (Third International Mathematics and Science Study) and PISA (Program for International Student Assessment). The TIMSS assessments are done in grades 4 and 8, and PISA at age 15. Results of 4 TIMSS and 3 PISA assessments are currently available, and 84 countries participated at least once in TIMSS (grade 8) or PISA. Several other international studies provide data points for additional countries.

Because TIMSS and PISA are by far the most reliable assessments, and adult proficiency is predicted better by proficiency in the 8th grade than in the 4th grade, we adopted the strategy of calculating the average of the 8th grade TIMSS scores and (age 15) PISA scores for countries participating in at least one assessment. For countries that did not participate in the 8th grade TIMSS or in PISA, missing data were extrapolated into this data set from the other assessments.

2.2. TIMSS and PISA

TIMSS assessments have been performed in a 4-year cycle, in 1995, 1999, 2003 and 2007. Tests of mathematics and science are administered in grades 4 and 8, generally with a larger number of countries participating in grade 8 than in grade 4. The results are publicly available at http://timss.bc.edu/timss2003.html and http://nces.ed.gov/timss/tables07.asp. Further information is available in Gonzalez, Galia, Arora, Erberber, and Diaconu (2004), Martin, Mullis, Gonzales, and Chrostowski (2004), Martin, Mullis, and Foy (2008), Mullis, Martin, Gonzales, and Chrostowski (2004), and Mullis, Martin, and Foy (2008).

PISA has been performed in a 3-year cycle, with results available from 2000, 2003 and 2006. Children aged 15 were tested in mathematics, science and reading. The reading test is defined as a measure of an individual's capacity to understand, use and reflect on written texts, and it is difficult to draw a distinction between this and intelligence. The results are available at http://www.oecd.org/dataoecd/30/18/39703566.pdf, http://nces.ed.gov/pubs2002/2002116.pdf, and http://pisacountry.acer.edu.au/. 74 countries participated in at least one TIMSS assessment, and 18 participated in all four; 57 countries participated at least once in PISA, and 30 participated in all three assessments.

In both TIMSS and PISA, the results are graded with methods based on item response theory, which models student proficiency as a latent variable. In both assessments the results are published separately for each tested subject and are reported on a 500/100 scale. In TIMSS the mean score of 500 is the average of those countries participating in the first TIMSS assessment in 1995, and in PISA it is the average of the participating OECD countries. In both TIMSS and PISA the individual-level, within-country standard deviation is about 85.

Within each assessment the scores of the different subjects were highly correlated at the country level, as expected from the results of earlier studies (Rindermann, 2006, Rindermann, 2007a, b). They were averaged separately for each of the four TIMSS and three PISA assessments. Minor trend adjustments were made by adjusting the means of the 18 countries that participated in all four TIMSS assessments to the same value. The same was done for PISA based on the 30 countries participating in all three assessments.

The averaged TIMSS scores (481.0 ± 50.0) and the averaged PISA scores (483.6 ± 51.9) were brought to the same mean and standard deviation (500 ± 50) for those 45 countries that participated in at least one TIMSS and one PISA assessment. These adjusted scores were averaged based on the number of assessments in which each country participated. Regressions in which the score was predicted by IQ and age at testing (which varied slightly among countries) showed no consistent age effect in either TIMSS or PISA.

These scores are a somewhat biased measure of intelligence because they measure only the proficiency of children who are still in school in grade 8 (TIMSS) or at age 15 (PISA). Large data sets for 8th-grade enrolment are not available, and therefore 8th-grade enrolment was estimated from data on youth literacy (YLit) and the proportion of children entering school who survive to grade 5 (Gr5), available from the Human Development Reports 2004 and 2007/08 (http://hdr.undp.org/en/reports/):

$$Enrolment = \left({}^1\!/_2 \times YLit + 0.5\right) \times Gr5^2.$$

The first term in this equation estimates the proportion of children entering school, and the second term estimates the proportion of those entering school who are still in school at age 15 or in grade 8. A conservative adjustment was made by assuming that those who are not in school would score half a (within-country) standard deviation (42.5 points) below those in school on scholastic assessment tests. The size of the schooling effect was estimated based on several studies on the effects of schooling on performance in school-related subjects (Alcock, Holding, Mung'ala-Odera, & Newton, 2008; Brouwers, Mishra, & Van de Vijver, 2006; Cliffordson & Gustafsson, 2008). Similar adjustments were made for other assessments that were done at different ages or grade levels.

2.3. Other assessments scored with methods of item response theory

Several assessments other than TIMSS and PISA were graded with modern methods of item response theory and published on a 500/100 scale. Those used for the extrapolation of data points missing in the original TIMSS/PISA data set are the following.

TIMSS 2007, 4th grade. This assessment included Kazakhstan and Yemen, which did not participate in any of the 8th-grade TIMSS or 15 year PISA assessments.

PIRLS Reading, 2001 and 2006. These are assessments of reading literacy of 4th-graders. 34 countries participated in 2001, and 39 in 2006. Data are available at http://nces.ed.gov/surveys/pirls/. These assessments provided data for Belize (2001) and Trinidad (2006).

IAE Reading 1991 assessed reading literacy of 9 and 14-year-olds in 30 countries. The results are published in Elley (1992). This assessment provided data for Trinidad and Venezuela at age 9 and 14, and Nigeria and Zimbabwe at age 14.

The raw scores were adjusted for age at testing in those assessments that showed non-trivial age effects. This was followed by adjustment for the approximate proportion in school at the age/grade of testing. To make the scores numerically equivalent to the TIMSS/PISA scale, the mean and standard deviation for each assessment were equalized with those of the TIMSS/PISA score for the countries participating in both kinds of assessment.

2.4. Older assessments

Some older assessments are available for which the results were published as "percent correct" scores.

IAEP Mathematics 1990/91. 19 countries participated in this assessment of 13-year-olds, of which China and Mozambique did not participate in TIMSS or PISA. Results are published in Lapointe (1992).

Second International Science Study 1983/84. Children from 23 countries were tested at age 14, and from 17 countries at age 10. The age 10 test provided data for Nigeria, and the age 14 test for China, Nigeria, Papua New Guinea and Zimbabwe. The results are published in Keeves (1992).

Second International Mathematics Study 1981. Mathematics was assessed in 13-year-olds. 17 countries participated, including Nigeria and Swaziland. The raw scores are published in Medrich and Griffith (1992).

First International Science Study 1970. Science was assessed at ages 10 (14 countries) and 14 (16 countries), including India. Results are published in Comber and Keeves (1973).

The results of these assessments show strongly nonlinear relationships with IQ and TIMSS–PISA score, and therefore nonlinear model fitting was employed before adjustments for age (if applicable) and proportion in school were made and mean and standard deviation were equalized with those of TIMSS–PISA.

2.5. The SACMEQ studies

The SACMEQ studies of 1995–1998 and 2000/01 are regional assessments of 6th grade reading and mathematics in the countries of South and East Africa. The results are available at http://www.sacmeq.org/indicators.htm. Data from the 2000/01 assessment are used except for Zimbabwe, which participated only in the 1995–1998 assessment. SACMEQ provides data for Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, Swaziland, Tanzania, Uganda, Zambia, Zanzibar and Zimbabwe. Results are published on a 500/100 scale.

Only two of the countries in SACMEQ (Botswana, South Africa) participated also in TIMSS, and none in PISA. For these two countries, the SACMEQ scores were 189 points higher than the TIMSS/PISA scores (weighted by the number of times they participated in TIMSS). SACMEQ scores for all participating countries were adjusted accordingly before an adjustment for proportion in school was applied. These adjusted SACMEQ scores correlated at $r = .678$ with measured IQs ($N = 11$ countries, $p = .022$) and $r = .655$ with IQs including estimates ($N = 13$ countries, $p = .015$).

2.6. Scaling to IQ metric

Two methods were employed to scale the 500/100 metric of the scholastic assessments to the 100/15 IQ metric. In one method, the average score of Britain was adjusted to 100 and the within-country standard deviation of 85 was adjusted to 15. These scores show the actual performance on the school assessment tests relative to Britain. The main difference between IQ and EA in this transformation is a 48.6% greater standard deviation for EA, based on the 86 countries that have scores for both. This means that relative to within-country differences, international differences are greater for EA than for IQ. In this sense, EA is more "culturally biased" than IQ. A likely reason is that the quality of schooling is generally lower in low-IQ countries, and this has a greater effect on school performance than on IQ (Lynn, Meisenberg, Mikk, & Williams, 2007). Conversely, the quality of schooling is generally higher in high-IQ countries, and this has a greater effect on school performance than on IQ.

Therefore a second scaling method was employed by adjusting the mean and standard deviation of EA to be Equal with IQ for those 86 countries with both scores. This makes EA and IQ more immediately comparable by showing whether a country's EA is higher or lower than expected based on that country's IQ.

EAs calculated by both scaling methods are included in Table 1, along with the measured IQs of those 86 countries that have data for both EA and IQ. The measured national IQs are those given by Lynn and Vanhanen (2006, p. 55–61), updated for some nations in Lynn (in press).

3. Validity of national IQs

3.1. Correlation between IQ and EA

The correlation (Pearson's r) between EA and IQ is .917 for the 86 countries that have both measured IQ and EA, and .907 for countries with measured and estimated IQs ($N = 108$). The corresponding nonparametric correlations (Spearman's ρ) are .917 and .905, respectively. These correlations are remarkably high. In Lynn and Vanhanen (2006), the correlation between different IQ studies in the same country was

Table 1

IQ and EA (educational attainment) for the 108 countries and provinces for which an EA score is available. Estimated IQs are in parentheses. EA is scaled in two ways: EAequ is calculated by equalizing mean and standard deviation of EA with IQ. This measure should be used when the question is whether a country's SA is higher or lower than expected from its IQ. EAact is the actual school performance of a country transformed to the IQ scale, with a mean for Britain of 100 and a within-country standard deviation of 15. Also given are scores for the quality of the IQ and EA data (IQqu and EAqu).

Country	IQ	IQqu	EAequ	EAact	EAqu
Algeria	(83)		84.4	78.4	2
Argentina	93	9	84.9	79.1	2
Armenia	92	3	94.1	92.7	4
Australia	98	10	100.1	101.8	14
Austria	100	4	99.5	100.8	8
Azerbaijan	(87)		86.8	82.0	2
Bahrain	81	2	88.1	83.8	4
Belgium	99	7	99.8	101.2	12
Belize	(84)		76.9	67.2	1
Bosnia	94	4	91.9	89.5	2
Botswana	71	2	79.4	71.0	4
Brazil	87	13	82.7	75.8	6
Bulgaria	93	6	94.7	93.7	10
Canada	99	6	100.6	102.5	12
Chile	90	9	86.9	82.1	6
China	105	16	100.8	102.7	2
Colombia	84	7	82.6	75.7	6
Cyprus	(91)		91.6	89.1	8
Croatia	99	7	96.4	96.2	2
Czech Republic	98	7	99.9	101.4	12
Denmark	98	5	96.7	96.7	8
Egypt	83	5	84.7	78.8	4
El Salvador	(80)		78.2	69.2	2
Estonia	99	7	101.0	103.0	4
Finland	99	5	102.8	105.8	8
France	98	9	98.4	99.1	8
Germany	99	16	98.1	98.7	8
Georgia	(94)		86.9	82.1	2
Ghana	71	6	68.1	54.2	4
Greece	92	10	94.1	92.8	8
Hong Kong	108	16	103.2	106.3	12
Hungary	97	8	99.3	100.5	14
Iceland	101	4	97.5	97.9	8
India	82	19	84.3	78.2	1
Indonesia	87	8	85.1	79.4	10
Iran	84	9	87.6	83.1	8
Ireland	92	10	99.5	100.9	8
Israel	95	14	94.3	93.1	10
Italy	97	13	95.1	94.3	14
Japan	105	20	103.5	106.8	14
Jordan	85	7	89.4	85.8	8
Kazakhstan	(94)		100.6	102.5	1
Kenya	72	11	79.4	71.0	1
Korea (South)	106	9	104.3	108.0	14
Kuwait	87	9	83.5	77.1	4
Kyrgyzstan	(90)		75.3	64.9	2
Latvia	(98)		95.8	95.3	12
Lebanon	82	4	86.9	82.1	4
Lesotho	(67)		66.5	51.7	1
Liechtenstein	(100)		100.1	101.7	6
Lithuania	91	5	95.7	95.2	10
Luxembourg	(100)		94.7	93.7	6
Macau	(101)		100.2	101.9	4
Macedonia	(91)		89.9	86.6	4
Malawi	60	3	62.3	45.5	1
Malaysia	92	7	96.0	95.6	6
Malta	97	2	93.7	92.3	2
Mauritius	89	5	81.3	73.7	1
Mexico	88	8	87.1	82.4	6
Moldova	(96)		91.6	89.0	4
Montenegro			86.8	82.0	2
Morocco	84	9	79.1	70.5	4

Table 1 (*continued*)

Country	IQ	IQqu	EAequ	EAact	EAqu
Mozambique	64	2	72.6	60.8	2
Namibia	72	2	66.2	51.3	1
Netherlands	100	10	101.1	103.1	10
New Zealand	99	11	98.6	99.5	12
Nigeria	69	6	72.4	60.5	4
Norway	100	2	96.5	96.4	12
Oman	85	7	84.8	78.9	2
Palestine	86	4	82.7	75.9	4
Papua New Guinea	83	4	92.0	89.6	1
Philippines	86	3	77.8	68.6	4
Poland	95	13	97.5	97.8	6
Portugal	95	6	93.6	92.0	8
Qatar	83	6	75.7	65.5	4
Romania	91	6	91.5	89.0	10
Russia	97	6	97.1	97.3	14
Saudi Arabia	80	6	80.5	72.6	4
Serbia and Montenegro	89	2	91.4	88.8	8
Seychelles	(86)		82.3	75.3	1
Singapore	108	5	107.7	113.0	8
Slovakia	96	4	98.4	99.1	10
Slovenia	96	7	99.4	100.7	10
South Africa	72	14	68.6	54.9	6
Spain	98	8	96.3	96.0	10
Swaziland	(68)		75.6	65.3	2
Sweden	99	8	98.7	99.6	12
Switzerland	101	6	99.7	101.0	8
Syria	79	7	87.2	82.6	2
Taiwan	105	18	104.9	108.8	8
Tanzania	72	8	77.0	67.4	1
Thailand	91	6	91.8	89.4	10
Trinidad	(85)		90.5	87.4	3
Tunisia	84	2	85.9	80.7	10
Turkey	90	8	89.5	86.0	8
Uganda	73	2	70.2	57.3	1
Ukraine	95	2	93.8	92.4	2
United Arab Emirates	83	6	93.7	92.2	2
United Kingdom	100	7	99.0	100.0	12
England			99.3	100.5	8
Scotland			96.3	96.1	6
USA	98	11	97.2	97.4	14
Uruguay	96	2	89.1	85.4	4
Venezuela	84	6	81.9	74.7	2
Yemen	83	6	67.5	53.3	1
Zambia	71	5	65.9	50.9	1
Zanzibar			70.3	57.5	1
Zimbabwe	72	4	72.8	61.1	3

given as .92 (p.62). For the school assessments, the correlation between TIMSS and PISA is .87 for those 45 countries that participated at least once in each study. Thus, the correlation between EA and IQ is not markedly different from correlations between two IQ studies administered in the same country, or two scholastic assessments administered in the same country. These correlations can be used as reliability coefficients to correct the correlations between EA and IQ for attenuation using the formula given by Ferguson (1971, p.370). With corrections for the unreliability (attenuation) of both IQ and EA, the correlation is 1.0.

Correlations can be markedly affected by range restriction. Therefore we will examine the relationship between IQ and EA by investigating the determinants of absolute differences between IQ and EA, rather than by investigating their correlations. For the 86 countries, the mean difference between EA (scaled by equalizing mean and standard deviation with IQ) and IQ is 3.27 points, and the median

difference is 2.29 points. The largest discrepancy is 15.46 points in Yemen, which has a far higher IQ than EA.

3.2. Data quality

If IQ and EA measure closely related or identical constructs, we can expect that the discrepancy between IQ and EA is lowest in countries with high quality of the IQ and EA data. To examine this, we constructed measures for data quality.

For IQ, a score for data quality was defined by adding two scores: (1) the number of IQ studies for a country and (2) a score for the total number of subjects that participated in all IQ studies available for the country. The latter was determined by giving a score of 1 if the total sample size was below 200, 2 if it was between 200 and 500, 3 between 500 and 1000, 4 between 1000 and 2000, 5 between 2000 and 5000, 6 between 5000 and 10,000, and a score of 7 was awarded for sample sizes above 10,000.

The quality of the scholastic assessments was estimated by giving countries 2 points for each grade 8/age 15 TIMSS or PISA assessment in which they participated. For countries that did not participate in TIMSS or PISA, 1 point was awarded for each of the other assessments. The maximum score was 14 for countries participating in all three PISA and all four TIMSS studies.

The discrepancy between IQ and EA correlated at $r = -.202$ with quality of the IQ data ($p = .062$), and $-.509$ with quality of the EA data ($p < .001$). This means that the higher the data quality, the lower is the discrepancy between IQ and EA.

These results suggest that limitations in data quality are more important for the scholastic assessments than for the IQ studies. Regression models were used to predict the discrepancy between IQ and EA jointly by quality of the IQ data, quality of the EA data, and either IQ, EA, or log-transformed GDP as a third variable. In all these models, only EA data quality emerged as a significant predictor, with p values of .001, .002 and <.001, respectively.

For countries with a score of 1–4 for EA data quality ($N = 37$), the average discrepancy between IQ and EA was 4.91 ± 3.34 points. For countries with an EA data quality score of 6 and higher ($N = 49$) it was 2.03 ± 1.55 points. The difference is highly significant ($p < .001$).

Closer examination of the major scholastic assessments shows that the average difference between two TIMSS assessments in the same country is the equivalent of 2.58 IQ points, and the difference between two PISA assessments is 1.61 points. However, the average discrepancy between a single TIMSS assessment and a single PISA assessment in the same country is 3.80 points. Although there is good agreement within each of the two study programs, the substantially greater discrepancy between TIMSS and PISA shows that this consistency is largely due to biases within each of the two study programs that are carried over from one assessment to the next. Discrepancies between two IQ studies that were done in the same country average 4.21 points, with a median of about 3 points.

3.3. Effects of kind of IQ test and test date

We can expect that highly reliable IQ tests and tests that include school-related tasks produce scores that are close to the results of scholastic assessments; and tests of fluid intelligence or other non-academic skills, as well as tests that are poor measures of anything, are more likely to produce large discrepancies. When IQ studies were investigated separately for the type of test employed, the following mean discrepancies between IQ and EA were found: 4.02 ± 3.37 points for the Standard Progressive Matrices (SPM: $N = 60$), 3.62 ± 3.71 for the Colored Progressive Matrices (CPM: $N = 35$), 2.19 ± 1.63 for the Wechsler Intelligence Scale for Children (WISC: $N = 16$), and 2.83 ± 2.61 for the Draw-a-Man test ($N = 10$). This analysis was restricted to the 73 countries with an EA quality score of 4 or more. The difference between the SPM and WISC is statistically significant at $p < .05$. The likely reason for this result is that the WISC but not the SPM includes measures of crystallized intelligence and school-related skills.

The TIMSS data range from 1995 to 2007, and PISA data from 2000 to 2006. The IQ studies, however, span many decades, and some are quite old. We know that national IQs can change substantially over time (Flynn, 1987), and therefore the IQ data in Lynn and Vanhanen (2006) and the present updates are corrected for the secular trend. If the Flynn effect corrections are accurate, and if the magnitude of the Flynn effect has been the same in all countries, the older IQ studies should be as useful as the more recent studies for the prediction of the EA. However, if the rate of the Flynn effect has been markedly different between countries, the more recent IQ studies are expected to be better than the old ones at predicting scholastic achievement.

The results show no relationship between the date of the IQ study and its match with the EA. 29 countries with an EA quality score of 4 or more also have IQ data for both before and after 1985. For these 29 countries the average discrepancy of individual IQ studies with EA is 3.47 points ($N = 67$ studies) for the old IQs, and 4.01 points ($N = 74$ studies) for the more recent IQs. The difference is in the "wrong" direction, but it does not even approach statistical significance.

3.4. Regional differences

To determine whether the relationship between IQ and EA applies equally to all parts of the world, EA and IQ were tabulated separately for 8 world regions: (1) Continental Europe (excluding ex-communist countries), (2) English-speaking countries with European-descended majorities, (3) ex-communist countries, (4) Latin America, (5) Middle East including North Africa, (6) sub-Saharan Africa, (7) East Asia, and (8) the rest of Asia (South Asia, Southeast Asia and Oceania).

Table 2 shows the results, including only countries with data for both IQ and EA. Comparisons of IQ with EA show that the discrepancies between IQ and EA are small in each world region. The largest discrepancy is found in Latin America, where the performance on scholastic tests is 3.9 points lower than predicted from IQ. Another noteworthy finding is that the East Asian countries do not over-perform on scholastic achievement tests. Their EA is actually 2.1 points lower than expected from their IQ.

Table 2 also shows that the country-level discrepancies between IQ and EA are smallest in the countries of continental Europe, and largest in Africa, the Middle East, and Asian

Table 2
Relationship between IQ and EA in 8 world regions. EA is calculated either by equalizing mean and standard deviation with IQ (EAequ), or as actual performance (EAact). Also shown is the average *absolute* difference between IQ and EAequ for the countries in each region (ΔIQ–EA).

Region	IQ	EAequ	EAact	ΔIQ–EA	N
Europe	98.3 ± 2.3	97.6 ± 2.7	98.0 ± 4.0	1.7 ± 1.2	16
English-speaking	97.7 ± 2.9	99.2 ± 1.2	100.3 ± 1.8	2.3 ± 2.7	6
Ex-communist	94.8 ± 3.1	96.1 ± 3.2	95.8 ± 4.7	2.1 ± 1.1	15
Latin America	88.9 ± 4.5	85.0 ± 2.8	79.3 ± 4.1	3.8 ± 2.8	7
Middle East	84.4 ± 3.8	84.8 ± 6.5	78.9 ± 9.7	4.6 ± 4.1	17
Africa	69.9 ± 3.9	71.3 ± 5.4	58.8 ± 8.1	4.7 ± 2.5	12
East Asia	106.2 ± 1.5	104.1 ± 2.3	107.6 ± 3.4	2.1 ± 2.0	6
Rest of Asia	87.1 ± 3.8	86.9 ± 6.5	82.1 ± 9.7	4.8 ± 3.4	7
All countries	89.8 ± 10.7	89.8 ± 10.7	86.4 ± 15.9	3.3 ± 2.9	86

countries other than those of East Asia and the Middle East. A likely reason is that data for IQ, EA or both tend to be more accurate in highly developed countries than in less developed countries.

Considering only the 86 countries with data for both IQ and EA, 91.0% of the variance for IQ and 83.6% of the variance for EA is between rather than within the world regions listed in Table 2. Therefore correlations between IQ and EA are expected to be low among countries within regions. Table 3 shows that with one exception the correlations within world regions are nevertheless positive, although many results are not statistically significant due to small sample size. In a regression model in which EA is predicted by the dummy-coded world regions and IQ, IQ is still a highly significant predictor ($t = 4.38$, $p < .001$, $N = 86$).

4. Discussion

We believe that the correlation of .917 between measured IQ and EA, attenuation corrected to 1.0, establishes without a reasonable doubt that national IQ is a valid measure for cognitive attainment at the country level. The correlation between these two variables is higher than the correlations that are commonly found between other "development indicators." For example, among the 86 countries with both IQ and EA, the correlation of log-transformed GDP is .753 with IQ and .734 with EA. For a composite measure of schooling that combines data about average length of schooling and school enrolment rates, these correlations are .768 and .738.

Table 3
Correlations between IQ and EA within world regions (Pearson's *r*), significance level (*p*) and sample size (*N*), separately for countries that have both measured IQ and EA and those having either measured or estimated IQ ("all IQs") along with EA.

Region	Measured IQs only			All IQs		
	r	p	N	r	p	N
Europe	.681	.004	16	.720	.001	19
English-speaking	−.101	.849	6	−.101	.849	6
Ex-communist	.789	<.001	15	.657	.001	22
Latin America	.812	.026	7	.637	.048	10
Middle East	.351	.167	17	.351	.154	18
Africa	.397	.201	12	.384	.175	14
East Asia	.511	.301	6	.710	.074	7
Rest of Asia	.374	.409	7	.374	.409	7

For Vanhanen's democracy index, the correlations are .659 and .619, respectively.

The high correlation between IQ and EA shows that these two measures are not merely two otherwise unrelated "development indicators." It rather shows that intelligence tests and scholastic achievement tests measure the same or nearly the same construct. To the extent that educational attainment is important for a country's economic or cultural destiny, IQ is important as well. We suggest that both can be used interchangeably as measures of "human capital."

It is nevertheless instructive to investigate the reasons for discrepancies between IQ and EA. One of the most unexpected results is that a measure for the quality of the IQ data has no independent effect on the discrepancy between IQ and EA. However, the quality of the EA data does have a substantial effect. Possibly our measure for the quality of the available IQ data is defective, but we can nevertheless conclude that an important contribution to IQ–EA discrepancies comes from the inaccuracy of the scholastic assessments, rather than the inaccuracy of IQ studies. It suggests a need to improve the representativeness of the samples that are tested in TIMSS, PISA and other international scholastic assessment programs.

A second reason for IQ–EA discrepancies is the widespread use of tests of fluid intelligence, especially the Progressive Matrices tests, as measures of intelligence. These tests appear to produce larger IQ–EA discrepancies than tests that also measure crystallized intelligence, such as the Wechsler. This finding, if it proves consistent in future studies, is not a flaw that needs to be remedied but a genuine difference in the abilities that are measured by scholastic achievement tests as opposed to IQ tests.

The non-relationship of the IQ–EA discrepancy with the date of the IQ study is puzzling. With few exceptions, the scholastic assessments that were computed into the EA date from 1995 and later. Therefore the EA should be more closely related to the results of recent IQ studies than to the results of older IQ studies. Failure to find this predicted relationship suggests that the Flynn effect has proceeded at a similar pace in most countries of the world, or at least in most of those countries for which both IQ and EA are available. The implication that the Flynn effect, which is well established in economically developed nations, has proceeded at a similar pace in most countries of the world has been confirmed in the economically developing nations of Brazil, Dominica and Sudan (Meisenberg, Lawless, Lambert, & Newton, 2005; Colom, Flores-Mendoza, & Abad, 2007; Khaleefa, Sulman, & Lynn, 2009).

References

Alcock, K. J., Holding, P. A., Mung'ala-Odera, V., & Newton, C. R. J. C. (2008). Constructing tests of cognitive abilities for schooled and unschooled children. *Journal of Cross-Cultural Psychology, 39*, 529–551.

Barnett, S. M., & Williams, W. (2004). National intelligence and the emperor's new clothes: A review of IQ and the Wealth of Nations. *Contemporary Psychology, 49*, 389–396.

Brouwers, S. A., Mishra, R. C., & Van de Vijver, F. J. R. (2006). Schooling and everyday cognitive development among Kharwar children in India: A natural experiment. *International Journal of Behavioral Development, 30*, 559–567.

Cliffordson, C., & Gustafsson, J. -E. (2008). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence, 36*, 143–152.

Colom, R., Flores-Mendoza, C., & Abad, F. J. (2007). Generational changes on the Draw-A-Man Test: A comparison of Brazilian urban and rural children tested in 1930, 2002, and 2004. *Journal of Biosocial Science*, *39*, 79–89.

Comber, L. C., & Keeves, J. (1973). *Science achievement in nineteen countries.* New York: John Wiley.

Elley, W. B. (1992). How in the world do students read? IEA study of reading literacy. *International Association for the Evaluation of Educational Achievement, Report No ISBN-92-9121-002-3* Accessed 15 June 2009 at: http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content/storage_01/0000019b/80/13/07/42.pdf

Ferguson, G. A. (1971). *Statistical analysis in psychology and education.* London: McGraw-Hill.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*, 171–191.

Gonzalez, E. J., Galia, J., Arora, A., Erberber, E., & Diaconu, D. (2004). Reporting student achievement in mathematics and science. In M. O. Martin, I. V. S. Mullis, & S. J. Chrostowski (Eds.), *TIMSS 2003 Technical Report* (pp. 274–307). Chestnut Hill (MA: TIMSS & PIRLS International Study Center Accessed 9 June 2009 at: http://timss.bc.edu/timss2003i/technicalD.html

Hunt, E., & Sternberg, R. J. (2006). Sorry, wrong numbers: An analysis of a study of a correlation between skin color and IQ. *Intelligence*, *34*, 131–139.

Irwing, P., & Lynn, R. (2006). The relation between childhood IQ and income in middle age. *Journal of Social, Political and Economic Studies*, *31*, 191–196.

Jencks, C. (1972). *InEAuality: A reassessment of the effects of family and schooling in America.* New York: Basic Books.

Kanazawa, S. (2006). Mind the gap… in intelligence: Re-examining the relationship between inEAuality and health. *British Journal of Health Psychology*, *11*, 623–642.

Keeves, J. P. (1992). Learning science in a changing world. Cross-national studies of science achievement: 1970 to 1984. *International Association for the Evaluation of Educational Achievement Report No. ISBN-92-9121-001-3* Accessed 14 June 2009 at: http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/15/2b/8d.pdf

Khaleefa, O., Sulman, A., & Lynn, R. (2009). An increase of WAIS-R IQ in Sudan, 1987-2007. *Journal of Biosocial Science*, *41*, 279–283.

Landes, D. S. (1999). *The wealth and poverty of nations: Why some are so rich and some so poor.* New York (NY): Norton.

Lapointe, A. E. (1992). Learning mathematics. *Educational Testing Service Report No. ETS-22-CAEP-01; ISBN-0-88685-120-3.* Princeton (NJ: ETS Accessed 16 June 2009 at: http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/24/21/ac.pdf

Lynn, R. (in press). National IQs updated. Mankind Quarterly.

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

Lynn, R., Meisenberg, G., Mikk, J., & Williams, A. (2007). National IQs predict differences in scholastic achievement in 67 countries. *Journal of Biosocial Science*, *39*, 861–874.

Lynn, R., & Vanhanen, T. (2006). *IQ and Global InEAuality.* Augusta, GA: Washington Summit Publishers.

Martin, M. O., Mullis, I. V. S., & Foy, P. (2008). TIMSS 2007 Science Report. Chestnut Hill (MA: TIMSS & PIRLS International Study Center Accessed 9 June 2009 at: http://timss.bc.edu/TIMSS2007/intl_reports.html

Martin, M. O., Mullis, I. V. S., Gonzales, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 international science report.* Boston, MA: IEA, Boston College.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of intelligence.* Baltimore: Williams and Wilkins.

Medrich, E. A., & Griffith, J. E. (1992). *International mathematics and science assessments: What have we learned? National Center for Education Statistics Research and Development Report.* Washington DC: U.S. Department of Education.

Meisenberg, G. (2004). Talent, character and the dimensions of national culture. *Mankind Quarterly*, *45*, 123–169.

Meisenberg, G. (2009). Wealth, intelligence, politics and global fertility differentials. *Journal of Biosocial Science*, *41*, 519–535.

Meisenberg, G., Lawless, E., Lambert, E., & Newton, A. (2005). The Flynn effect in the Caribbean: Generational change of cognitive test performance in Dominica. *Mankind Quarterly*, *46*, 29–69.

Morse, S. (2008). The geography of tyranny and despair: development indicators and the hypothesis of genetic inevitability of national inEAuality. *Geographical Journal*, *174*, 195–206.

Mullis, I. V. S., Martin, M. O., & Foy, P. (2008). TIMSS 2007 international mathematics report. Chestnut Hill (MA: TIMSS & PIRLS International Study Center Accessed 9 June 2009 at: http://timss.bc.edu/TIMSS2007/intl_reports.html

Mullis, I. V. S., Martin, M. O., Gonzales, E. J., & Chrostowski, S. J. (2004). *TIMSS 2003 international mathematics report.* Boston, MA: IEA, Boston College.

Rindermann, H. (2006). Was messen internationale Schulleistungsstudien? Schulleistungen, Schülerfähigkeiten, kognitive Fähigkeiten, Wissen oder allgemeine Intelligenz? [What do international school achievement studies measure? School achievement, pupils' abilities, mental abilities, knowledge or general intelligence?]. *Psychologische Rundschau*, *57*(2), 69–86.

Rindermann, H. (2007). The big/G/-factor of national cognitive ability (author's response on open peer commentary). *European Journal of Personality*, *21*, 767–787.

Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: the homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality*, *21*, 667–706.

Rindermann, H., & Ceci, S. J. (2009). Educational policy and country outcomes in international cognitive competence studies. *Perspectives in Psychological Science*, *4*, 551–577.

Rindermann, H., & Meisenberg, G. (2009). Relevance of education and intelligence at the national level for health: The case of HIV and AIDS. *Intelligence*, *37*, 383–395.

Rushton, J. P., & Templer, D. I. (2009). National differences in intelligence, crime, income, and skin color. *Intelligence*, *37*, 341–346.

Smith, A. (1776). Inquiry into the Nature and Causes of the Wealth of Nations. Chicago: University of Chicago Press.

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

Volken, T. (2003). IQ and the wealth of nations: a critique of Richard Lynn and Tatu Vanhanen's recent book. *European Sociological Review*, *19*, 411–412.

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

Further reading

Beaton, A. E., Martin, M. O., Mullis, I. V. S., Gonzales, E. J., Smith, T. A., & Kelly, D. L. (1996). *Science achievement in the middle school years.* Chestnut Hill, MA: Boston College, TIMSS International Study Center.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzales, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years.* Chestnut Hill, MA: Boston College, TIMSS International Study Center.

Husen, T. (1967). *International achievement in mathematics: A comparison of twelve countries.* New York: John Wiley.