What Makes Nations Intelligent?
Earl Hunt

*Perspectives on Psychological Science* 2012 7: 284
DOI: 10.1177/1745691612442905

The online version of this article can be found at:
http://pps.sagepub.com/content/7/3/284
What Makes Nations Intelligent?

Earl Hunt
Department of Psychology, University of Washington, Seattle

Abstract
Modern society is driven by the use of cognitive artifacts: physical instruments or styles of reasoning that amplify our ability to think. The artifacts range from writing systems to computers. In everyday life, a person demonstrates intelligence by showing skill in using these artifacts. Intelligence tests and their surrogates force examinees to exhibit some of these skills but not others. This is why test scores correlate substantially but not perfectly with a variety of measures of socioeconomic success. The same thing is true at the international level. Nations can be evaluated by the extent to which their citizens score well on cognitive tests, including both avowed intelligence tests and a variety of tests of academic achievement. The resulting scores are substantially correlated with various indices of national wealth, health, environmental quality, and schooling and with a vaguer variable, social commitment to innovation. These environmental variables are suggested as causes of the differences in general cognitive skills between national populations. It is conceivable that differences in gene pools also contribute to international and, within nations, group differences in cognitive skills, but at present it is impossible to evaluate the extent of genetic influences.

Keywords
behavioral economics, education, cognition, individual differences

Are some nations smarter than others? If they are, does it make any difference? The gist of this article is that the answer to both questions is “yes.” The argument is presented in five sections. First, some data will be presented showing how heavily modern society depends on cognitive skills. Second, there will be a presentation of a model of the development of individual cognitive skills, emphasizing the importance of these skills in the industrial and postindustrial societies. The third section discusses the measurement of national intelligence and correlates of those measurements. The fourth discusses genetic and environmental influences on national supplies of cognitive skills. The fifth and final section discusses conceptual issues associated with the idea of national cognitive skills.

Before beginning the substantive discussion, a brief word about terminology is in order. From a rhetorical perspective, the simplest way to present this article would be to discuss national intelligence. However, the word “intelligence” has acquired a great deal of rhetorical baggage. To some, the word brings up images of genetically determined cognitive power. Such an interpretation would raise issues that are not germane to the discussion here. Therefore, in order to avoid unnecessary controversy, I will follow the example of Schmidt and Hunter (1998) and use the term general cognitive ability (GCA). This is meant to refer to the general factor associated with, and revealed in performance on, a variety of assessments of cognitive skills. These assessments include many avowedly educational exams, where educational achievement is associated with GCA.

A discussion of national differences obviously has to have terms to refer to groups of nations. A variety of terms are in common use, including “developed,” “developing,” “first [second, third] world,” and “modern versus traditional” societies. All these terms have problems. Probably the most acceptable is the “developed”–“developing” distinction, largely because these are the terms used by many international agencies to classify nations. Thus, Switzerland is a developed country, while Senegal is a developing country. This contrast is not suitable for the present purposes, for it implies that there is an arc of development that nations should, or perhaps must, follow in order to become “developed.” (At some future time, will Senegal be just like Switzerland? If there is no such time, does this mean that Senegal will forever have a gap in its development?) To avoid this issue, I will use terms that, insofar as possible, describe the dominant features of a country’s socioeconomic structure. In particular, countries such as the United States, Canada, Europe, and Japan will be referred to as industrial and postindustrial countries (I&PI). In most cases, all that will be needed is a distinction between I&PI and non-I&PI countries (again, like Senegal). It is important to remember that the I&PI versus non-I&PI distinction is a fuzzy one. Many countries (e.g., Chile) are moving toward I&PI society.
but still have substantial portions of their populations outside of the I&PI system.

On occasion, it will be appropriate to refer to countries by region, such as the sub-Saharan African or Latin American countries, or by some dominant feature of their society, for example, Islamic countries.

### Our Cognitive Society

Archaeologists identify periods of hominid development by the tools that the hominids used. Australopithecines made the crude Oldowan stone tools 2.5 million years ago. The Paleo-Indians of North America fashioned the elegant stone Clovis spear points just 12,000 to 14,000 years ago. For most of hominid history, tools were developed in order to extend human physical capacities, making it possible to break open oyster shells with a hand axe and then, millennia later, hunt bison with a bow and arrow. Tools to develop physical capacities are still being developed. Today we can fly.

About 10,000 years ago, a radically new type of tool appeared, the cognitive artifact. Rudimentary number systems, and then techniques of writing, were developed (Wolf, 2008). Number systems led to logical reasoning. Writing made it possible to communicate ideas across space and (one way) through time. The King James Bible, written in the 17th century, contains information collected from original writings in the first century of the current era, strained through the social and political thinking of England’s Jacobian era. The King James Bible is still being read in the 21st century (Nicolson, 2005).

Cognitive artifacts were being shaped by society, and were shaping society, long before the King James Bible. A Sumerian clay scroll dating from roughly 3000 BCE said:

The day the import of your tablet was made known to me, I provided your agent with three minas of silver for the purchase of lead. Now, if you are still my brother, let me have my money by courier. (James & Thorp, 1994, p. 521)

The scroll was an invoice. On its face, the scroll revealed a society that had three extremely important cognitive artifacts: literacy, arithmetic, and a monetary system. Indirectly it implied three other things: a society that communicated at a distance, that could develop social contracts, and in which people had a sense of honor and obligation, at least between those who saw themselves as spiritual brothers.

Cognitive artifacts fall into two broad classes. **Physical artifacts**, such as the scroll itself and its inscribed writing system, are physical devices that can be used to extend cognitive and perceptual functions. **Mental artifacts**, such as logic and the scientific method of reasoning, are ways of thinking that are used to reason effectively about the phenomena we observe. Table 1 lists just a few of our modern cognitive artifacts. The list includes physical artifacts, like computers, and mental artifacts, such as logic, that bring order to our thinking. In order to be successful in our society, a person has to have the abilities required to use these artifacts.

In theory, a society that contains specialized cognitive tools might offer ecological niches to people with a variety of cognitive skills. Instead, however, modern society appears to demand a GCA that can then be shaped into specialized skills by training. This can be shown by an analysis of the cognitive requirements of different occupations.

The Department of Labor maintains a database, O*NET, that contains ratings of the cognitive skills required in over 850 occupations, ranging from **physicist** to **clothing model**. Jobholders within these occupations constitute over 90% of the American workforce. Occupations are rated for 20 different cognitive skills, ranging from hand–eye coordination to the ability to understand both spoken and written language. Hunt and Madhyastha (2012) found that variation in the need for these skills, across occupations, could be represented in a space defined by three orthogonal factors. The largest factor, which accounted for 45% of the generalized variance in the ratings, was a GCA factor. It had substantial loadings on almost every one of the skill ratings, with the highest loadings being on tasks involving logical reasoning and/or language comprehension. The American workplace evidently demands a high level of general ability, but with a strong verbal flavor. The second factor, perceptual and attentional skills, accounted for 19% of the variance, and the third factor, numerical skills, accounted for only 10%.

If GCA is in such high demand, one would expect the workplace to pay for it. This is so. Median incomes, across occupations, are exponentially related to the level of GCA required by the occupation ($R^2 = .60$).\(^1\)

With a very few exceptions, people who want a rewarding place in a developed society have to bring intelligence both to the workplace and to their daily lives. Gottfredson (1997) made this point elegantly, by arguing that so much of our everyday activities require cognitive guidance that daily life is, in a way, an intelligence test. The difficulty of this test has increased as society has become more complex, interdependent, and technologically oriented. A century ago, consumer

---

### Table 1. An Incomplete List of the Cognitive Artifacts Central to Modern Society

<table>
<thead>
<tr>
<th>Physical cognitive artifacts</th>
<th>Mental cognitive artifacts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written documents</td>
<td>Literacy</td>
</tr>
<tr>
<td>Maps</td>
<td>Mathematics</td>
</tr>
<tr>
<td>Printing</td>
<td>Law</td>
</tr>
<tr>
<td>Telecommunication</td>
<td>Logic</td>
</tr>
<tr>
<td>Computers and electronics</td>
<td>Scientific reasoning</td>
</tr>
<tr>
<td>Global positioning via satellites</td>
<td>Formal systems of finance and banking</td>
</tr>
</tbody>
</table>

Note. Cognitive artifacts are listed in their historical order of origin.
credit management was an easily solved problem. Unless you were wealthy, you paid cash. Today there are a bewildering number of choices, and paradoxically, the tools for managing personal wealth are more likely to be available to the well-to-do than to the poor (Bertrand, Mullainathan, & Shafir, 2006). Similarly, within developed societies, health management is increasingly dependent on a person’s understanding of medical advice, both to avoid problems and to manage illnesses. Such understanding is directly related to literacy skills, which are in turn an indicator of general intelligence (Gottfredson, 2004). In order to be healthy and wealthy, it helps to be wise.

And how do we become wise? GCA is used directly, to solve problems faced in the workplace and in daily life, and indirectly, to benefit from training required to develop specialized skills. Today we see specialization in the extreme. To illustrate, in Hunt and Madhyastha’s (2012) data set, two of the occupations with the highest GCA factor scores were physicists and surgeons. Obviously physicists and surgeons know different things and have different skills. In both cases, the acquisition of the necessary knowledge and skill depends on possession of a high level of GCA. Interactions between personal interests, personality, and opportunity will determine how GCA is shaped into usable cognitive skills.

The relation between GCA and training is particularly important, for education pays. In 2010, during a substantial recession, adults over 25 who were employed full time and held doctorates and professional degrees earned approximately 2.5 times more than fully employed high school graduates. Even more important, the degree holders were five times more likely to be employed than were the high school graduates.2

The Conceptualization of General Cognitive Ability

The conceptualization used here

The position taken here is that an individual displays GCA by solving the problems presented by his or her society. In I&PI societies, such problem solving typically involves use of both physical and mental cognitive artifacts. This conceptualization of cognitive ability was initially presented as part of collaborative work with Jerry Carlson (Hunt & Carlson, 2007) and is further elaborated upon in Hunt (2011). It differs from the view that the primary role of a theory of GCA is to explain individual differences in test scores. However, the evidence gathered by studying test scores is regarded as relevant to the understanding of both individual and national GCA. My position is close to some of the views expressed in Robert Sternberg’s many writings on the importance of “practical intelligence” (e.g., Sternberg et al., 2000) and to Philip Ackerman’s emphasis on knowledge as being the “dark matter” of adult intelligence that is unevaluated by conventional tests (Ackerman, 2000). The two features that I believe are unique are the emphasis on the use of cognitive artifacts as being a central feature of GCA and the argument that the concept of cognitive ability can usefully be generalized to apply to large social groups, rather than being confined to individuals.

History provides us with an example of both the use of cognitive artifacts and the idea that societies can differ in their level of GCA. When the “Black Death” (bubonic plague) outbreak occurred in the 14th century, the disease was variously blamed on the sinful acts of the population and foul air emanating from the corpses of the victims. Among the actions taken were prayer, the granting of indulgences, and vigorous programs of corpse burning. Today we treat bubonic plague as a bacterial infection and have identified the flea as a vector of the disease. Our methods of prevention rely on cognitive artifacts, ranging from the germ theory of disease (a mental cognitive artifact) to the use of the media (physical cognitive artifacts) to spread information about disease control. In addition to specialized knowledge, the modern approach relies on GCAs, including ways of thinking about disease, which were not available to medieval society.3

It follows that an individual’s level of GCA is the result of a developmental process in which a person acquires expertise in the use of the cognitive artifacts available in his or her society (Hunt, 2011). This view has precedents, especially in the work of Stephen Ceci (1990). Once again, though, the model proposed here differs from earlier ones in its emphasis on the role of cognitive artifacts. This point is clear when we consider how intelligence is developed over time.

The developmental process is outlined in Figure 1A. The panel shows three classes of influences on cognitive ability: genetics, the physical environment, and the social environment. A naive view of intelligence is that it is a constructed process, rather like building a computer. The computer builder begins with a plan, uses that plan to assemble devices available in the physical environment (once a complex of circuits and vacuum tubes, now an array of silicon chips and information storage devices), and when the physical computer is developed, it is exposed to the social environment (software), to create a device that, at least by some definitions, can think. The process of giving cognitive abilities to a computer can serve as a first approximation to the development of GCA in people. It is captured in the solid arrows in the panel.

Cognitive ability is not inherited directly. The genetic potential for acquiring cognitive ability is. There is ample evidence that individuals differ in their genetic potentials. That potential is then developed, in two ways. The first development depends on the physical environment, which includes such things as nutrition and the extent of exposure to infectious disease. The physical environment influences the brain processes and functions that provide the individual with information processing capacities, such as the working memory–attention complex. These are not cognitive abilities in themselves, any more than a carpenter’s tools are equivalent to a carpenter’s skills. Cognitive abilities are acquired when the brain is filled with the knowledge and techniques of reasoning that are known to the person’s society.
The computer analogy is accurate as far as it goes, but it does not provide for the feedback loops that are characteristic of biological and ecological systems. These are shown by the dashed lines in Figure 1A. Genetic processes modulate an organism’s biological processes throughout life. Many of these continuing genetic actions are triggered by properties of the physical environment (Johnson, Penke, & Spinath, 2011). The physical environment is obviously influenced by the social environment. Finally, the individual can exert considerable influence on his or her environment. The extent to which this is possible varies considerably across societies. However, some influence is always possible.

Societies change, and with social change comes change in GCA. In the 20th century, cognitive test scores rose at a far faster rate than could be accounted for by genetic selection (Flynn, 2007). The cognitive artifacts that we are developing today ensure that this process will continue in the future. To take just one example, the development of the World Wide Web and, along with it, powerful search engines has made more information available to individuals than ever before in history. However, it requires considerable skill to be a sophisticated user of this information. Equally important, collaborative thought is being facilitated by first the Internet and now the various social media built on it. It is interesting to speculate what sorts of collaborations might have developed between the great French chemist Antoine Lavoisier (1743–1794) and the shrewd American experimenter Benjamin Franklin (1706–1790) if they had only known each other’s e-mail addresses.

In an important sense, we have more cognitive abilities than our predecessors had. By using our cognitive artifacts, we are able to solve problems that they could not. Barring a social catastrophe, such as a worldwide nuclear war, it is inevitable that a substantial fraction of our descendants will be more intelligent than we are, for they will have better cognitive artifacts than we have. However, there is no guarantee that the benefits of these artifacts will be equally distributed across the globe. Before developing this point, though, an aside about the measurement of cognitive ability is in order.
Conventional cognitive testing

The conceptualization of GCA presented here does not imply a rejection of all previous approaches to intelligence. The various analyses of intelligence offered since testing was developed have resulted in valuable information. It would be foolish to disregard these results. It is more constructive to consider the strengths and weaknesses of the approach that has dominated the study of individual differences in cognition since the introduction of intelligence testing by Alfred Binet early in the 20th century.

What Binet showed, and what has been shown over and over again in the hundred years following his work, is that some important general cognitive abilities can be evaluated in a testing situation in which an examiner, usually a stranger, evaluates a person’s cognitive capacities within a period of approximately 1 to 3 hr, in a setting divorced from the examinee’s normal life. We may think of this, only somewhat facetiously, as the “drop in from the sky” testing paradigm.4

“Drop in from the sky” examinations can be used to evaluate an examinee’s ability to think rapidly about problems, to develop a rough picture of the examinee’s working memory capacities, and to obtain a valid sample of the examinee’s knowledge about society. Such examinations include rudimentary testing of the examinee’s ability to comprehend language and to use simple mathematics. What the testing paradigm cannot encompass are evaluations of those cognitive skills that allow people to deal with issues over time. A particularly important set of traits that are vital to cognition but outside the testing paradigm are the traits related to self-organization, scheduling activities, establishing priorities, and, what may be more important, keeping those priorities in mind during future activity (Duckworth, Peterson, Matthews, & Kelly, 2007).

This deficiency of “drop in from the sky” testing is dramatically illustrated by the case of Henry Moulson (H. M.), which has been described in many introductory texts. H. M. suffered damage to his hippocampus, which made it impossible for him to store new episodic memories. As a result, he had to reside in a highly assisted living situation from the time of his injury, when he was in his 20s, until he died, in his 70s. Following the injury, his Wechsler Adult Intelligence Test scores were well within the normal range. As a result, H. M. is often described as having suffered profound retrograde amnesia with unchanged intelligence. In my view, it staggers the mind to say that a man who has to live in an assisted living situation because of a cognitive deficit has “normal intelligence.” One of the things that H. M. and similar hippocampal cases show is that tests within the conventional framework provide a limited evaluation of the GCA skills required in human society.

In summary, conventional tests of intelligence, aptitude, and educational achievement provide a partial evaluation of a person’s GCA, but they certainly do not provide a complete picture of it. However, saying that an evaluation is incomplete is not the same as saying that it is useless.

How useful are the tests?

The tests that have been developed to fit the “drop in from the sky” paradigm vary considerably in format and are often carefully not labeled “intelligence tests” (Hunt, 2011, Chapter 2). Nevertheless, the correlations between tests are so high that, at the level of generality of the present discussion, they can all be treated together.

Many reviews have been written describing the efficacy of the tests as personnel selection devices in academic, industrial, or military settings. Three findings stand out. One is that the tests are not perfect predictors of performance. Although examples of failure of prediction are often put forward as arguments against testing itself, such attacks are directed at a straw person. The most enthusiastic advocate of testing has never claimed a validity coefficient of .9. The second finding is that the tests do a better job of predicting performance in both academic and school settings than any other screening device of comparable cost and that they certainly do much better than personality tests.5 The third is that the important statistic is not the correlation between test scores and the performance of students or jobholders who have been selected; it is the predictive correlation, the estimated correlation between test score and job performance in the applicant population.6 Many of the attacks on testing have failed to consider this point.

There are three sources of evidence for the efficacy of tests within a developed society. The first is an analysis of the results from the many studies that have been conducted of the use of test scores to predict performance within the normal range of test scores. Examples include studies of the relation between entry test performance and performance in colleges and universities, which cover roughly the top half to two-thirds of the presumed range of intelligence in the population, and the smaller, but still substantial, studies that have been done of the performance of people with unusually high scores or scores that are in the low to normal range.

Studies of individuals in the normal range of test scores.

Kuncel and Hezlett (2010) conducted a meta-analysis in which they estimated predictive correlations between test scores and performance in academic (college/university) and workplace settings. Their results are summarized in Table 2. These correlations show that in typical industrial and educational settings, increasing the test performance of successful applicants by one standard deviation unit should result in improvement in selected applicants’ work or academic performance by slightly more than half a standard deviation unit. The plural “applicants” is important. As a general rule, people with high test scores will perform better than people with low test scores, but exceptions are to be expected.

Kuncel and Hezlett’s (2010) estimates apply to populations in the “average” range of cognitive ability, across both the workplace and academia. The following results deal with people with “low average” and “very high” test scores.
Low average test scores. In the United States, prospective military recruits are given a cognitive test battery called the Armed Forces Qualifying Test (AFQT). AFQT scores are used to screen potential enlistees into five categories, with I being the highest and V the lowest. People who score in Category V are not recruited. Category IV scores are those from the 10th to 30th percentile of the norming population. Recruitment from this category is limited by law.

The performance of Category IV soldiers was studied in detail during the 1960s and 1970s. Although the majority of the Category IV enlistees completed their first enlistment, first-year discharge rates were from two to three times higher than the discharge rates of others in the same service. Those who remained in the military were not promoted as rapidly as servicemen (there were no women) in a control group. In both groups, the modal rank achieved after 13 years was staff sergeant in the army or its equivalent in the other services. However, less than 20% of the Category IV soldiers held ranks above staff sergeant, while just under 40% of the control group were either first sergeants or master sergeants (or equivalent ranks). First and master sergeants are considered senior non-commissioned officers and usually have positions of substantial responsibility.

The authors of the study concluded that “the lower aptitude groups are 80–90 percent as effective as average-aptitude personnel” (Sticht, Armstrong, Hickey, & Caylor, 1987, p. 191). Many of the Category IV men served their country honorably and well. Nonetheless, as a group, their performance was reliably lower than that of servicemen with higher test scores.

Very high test scores. Louis Terman’s lifelong study of “genius” (his term for people with very high intelligence test scores) is one of the best-known studies in psychology. As a group, the people Terman studied had quite successful lives (Terman & Oden, 1959). Terman pointed out that the particular types of success enjoyed by his participants had been very much influenced by the historic events of their times: primarily the Great Depression and World War II. High test scorers also fare well in contemporary American society.

The Study of Mathematically Precocious Youth (SMPY) is a very large longitudinal study of people who were selected on the basis of test scores earned in their early teens. It has been estimated that the participants in this study are in anywhere from the top 1% to the top 0.01% of the U.S. population, the top one in 10,000, in terms of SAT test scores. One way to compare success between groups within the SMPY is to use the odds ratio, which is defined as

\[
\text{Odds Ratio} = \frac{\text{Percent of people achieving criterion in the target group}}{\text{Percent of people achieving criterion in the reference group}}
\]

Table 3 shows odds ratios when the target group is SMPY participants estimated to be in the 99.75th percentile in terms of all 13-year-olds, and the reference group is those “only” in the 99th to 99.25th percentile. Odds ratios are shown for a variety of social accomplishments, calculated just before the participants reached age 40. This was slightly more than 25 years after they had taken the test. Recall that these odds ratios were calculated within the top 1% of test scorers. The contrast would be even greater if we considered less selective reference groups. For example, by age 40, about 12% of the SMPY target group held patents. The base rate for patent holding in the U.S. population is 1%, which includes all age ranges. Using the population as a whole, the odds ratio for the top scorers would be 12:1. Even this is an underestimate, for slightly more than half the patents awarded annually are to people over 40 years of age (Feyrer, 2008), older than the people in the SMPY at the time Lubinski calculated the ratios.

Lubinski (2009, p. 358) concluded that

The likelihood of exceptional achievement is markedly enhanced as a function of general ability. There does not appear to be an “ability threshold” (i.e., a point at which, say beyond an IQ of 115 or 120, more ability does not matter).

Table 3. Odds Ratios Comparing Accomplishments of SMPY Participants With SAT Scores in the Fourth and First Quartiles of the Top 1%

<table>
<thead>
<tr>
<th>Achievement</th>
<th>Odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achieving the doctorate (including MD, LLD)</td>
<td>3.56</td>
</tr>
<tr>
<td>Publication in science, technology engineering, or mathematics</td>
<td>4.97</td>
</tr>
<tr>
<td>Holding one or more patents</td>
<td>3.01</td>
</tr>
<tr>
<td>Income in the top 95th percentile nationally</td>
<td>2.31</td>
</tr>
<tr>
<td>Literary publications</td>
<td>4.55</td>
</tr>
</tbody>
</table>

Note. Data are from Lubinski (2009). SAT scores were recorded when participants were 13 years old. Their achievements were tallied when they were 40 years old. SMPY = Study of Mathematically Precocious Youth.
This is essentially the same conclusion that was reached by Terman 50 years earlier (Terman & Oden, 1959). The finding has generalized over time and massive social changes.

Highly placed academics (Muller et al., 2005) and important media figures (Brooks, 2011; Gladwell, 2008) have maintained that very high scores do not count. They do.7

Conclusions concerning the intelligence of individuals

The argument to this point leads to four conclusions:

1. The industrial–postindustrial workplace (the I&PI society) requires a workforce with general cognitive abilities that can be trained for a variety of specialized occupations.
2. Intelligence should be conceptualized as skill in using the cognitive artifacts provided by society in order to solve important problems that the individual encounters.
3. The conventional “drop in from the sky” method of testing captures some, and evaluates some, but not all of the important GCAs required by I&PI societies.
4. In spite of their limitations, conventional intelligence tests and their analogs, such as the SAT and AFQT, capture enough individual variation in intelligence that the test scores are nontrivial predictors of cognitive achievements within our society.

The next step is to move from a focus on the intelligence of individuals to a focus on the intelligence of groups.

Are Nations Intelligent?

Is it meaningful to say that nations are intelligent? A case study is instructive.

Modern Japan, a country with very few natural resources, is one of the economic powerhouses of the world. Although Japan’s growth rate slowed somewhat during the 2000–2010 period, its economy was still the third largest in the world. In 2010 Japan’s gross domestic product per capita, corrected for local purchasing power (GDP-PPP/c), was $32,000. This value is comparable to that of France, Germany, and the United Kingdom. As of 2010, the life expectancy at birth of a Japanese citizen was 82.2 years. The infant mortality rate was 2.8 per thousand births. One hundred percent of the Japanese population had access to reliable water supplies and sanitary facilities. By any standard, Japan is an I&PI nation.

By contrast, modern Nigeria is one of the poorer nations of the world. Nigeria used to be an exporter of food; today it is an importer. In spite of being a supplier of oil, which is arguably the world’s most sought-after resource, Nigeria’s 2010 GDP-PPP/c was $2,500, less than a tenth of Japan’s. In 2010, a Nigerian citizen’s life expectancy at birth was 47.5 years, and the infant mortality rate was a staggering 97.4 per thousand births. Only 58% of the Nigerian population had access to reliable water, and only 38% had access to adequate sanitary facilities. In spite of its substantial physical resources, Nigeria is not an I&PI nation.8

The cognitive artifacts of Japan—well-organized schools, a commitment to the use of the most modern technologies, and a legendary work ethic—act to multiply the nation’s physical resources. In Nigeria, these cognitive artifacts are in short supply. There are many reasons why this might be true. Delving into them would not be germane to this article. The point is that, for whatever reason, Japanese society contains a greater supply of GCA than does Nigerian society.

There are two sorts of data used to measure GCAs of national populations. One is a compilation of scores on avowed intelligence tests, such as the Wechsler tests or the Raven Matrix tests, taken by residents of various countries. Richard Lynn and Tutu Vanhanen (2002, 2006) deserve credit for initiating a line of research based on this source of data. The line has since been pursued vigorously. In their various reports, they used data from somewhat different selections of countries. Regardless of exactly which data set was used, the correlation between national test scores and measures of GDP-PPP/c, possibly the best single index of economic strength, was about .65.

Within I&PI nations, the scores Lynn and Vanhanen used are likely to be accurate representations of a population’s cognitive skills. Considerable care was taken to make the Spanish, British, and American versions of the Wechsler Adult Intelligence Test comparable and to obtain validation samples that were representative of the country for which the test was being validated. The use of test scores outside of I&PI nations has been criticized on three grounds. One is that although the tests evaluate skills required by the I&PI society, they may not reflect the skills required by various societies outside that group. The second is that testing itself is so foreign to societies outside the I&PI world that the resulting scores are not valid. The third is that the test scores are often based on small and unrepresentative samples, rather than samples reflecting the population of the nation under study.

For the present purposes, the first objection is not crucial. If our purpose is to determine whether a nation has the cognitive resources appropriate for participation in the international economy, then the appropriate tests are those that evaluate the skills required to utilize the cognitive artifacts vital to that economy. The skills associated with orienting oneself in a jungle or desert may be vital in a traditional society, but in a modern industrial society, the skills needed to read a street map or use a global positioning system are more useful. When comparing two countries, then, the test scores (and statistics developed from them) are valid as measures of the extent to which the countries are prepared for the I&PI world. Whether the test scores are valid measures of the cognitive abilities required to participate in other societies, such as an agricultural, nomadic, or hunter-gatherer society, is an open question and can be answered only on a case-by-case basis.
The second objection, that the “drop in from the sky” paradigm is foreign to societies other than I&PI countries, also can be answered only on a case-by-case basis. Nations are not “in” or “out” of the I&PI world; different nations are in or out to some degree. To clarify, although there are certainly regions in the countries outside of the I&PI world where the use of tests would be a foreign practice, some form of testing is associated with formal education systems and these are widespread throughout the world. (China has been using formal examinations for educational screening for over a millennium.) Indeed, if the testing paradigm were foreign to a large proportion of the population, this alone would be evidence that a particular nation had a weak formal education system and, along with that, lacked many of the GCAs required in the I&PI world.

The third objection, that in some nations the test scores are drawn from atypical samples, is relevant here. In fact, it is especially relevant if the scores offered for a nation as a whole include data from an unrepresentative group for whom the test paradigm is not appropriate. Lynn and Vanhanen have been sharply criticized for including such data in their calculations, especially those reported for sub-Saharan Africa (Hunt, 2011, Chapter 11; Wicherts, Dolan, Carlson, & Van der Maas, 2010a, 2010b; Wicherts, Dolan, & Van der Maas, 2009, 2010). It is likely that the selection methods used by Lynn and Vanhanen resulted in spuriously low estimates of GCA in the nations outside the I&PI world. However, there is evidence that the biases are not sufficient to make a major alteration in the correlations between test scores and measures of national well-being, such as GDP-PPP/c (Whetzel & McDaniel, 2006).

An alternative way to measure national levels of cognitive competence is to use internationally accepted evaluations of educational progress, such as the Program of International Student Assessment (PISA). The educational data are typically comparable and the sampling somewhat representative across countries. However, the results apply to the contemporary student population, not to the general population. Educational data are available for virtually all I&PI nations and for many outside of the I&PI world. However, educational data are usually not available for those nations furthest from the I&PI world. For instance, no sub-Saharan country participated in the 2009 PISA studies, nor did any of the poorer nations of Latin America. In addition, the international assessment programs focus heavily on only a few topics. Educational fields relative to science, technology, engineering, and mathematics (STEM) are given special emphasis because the STEM fields are believed to drive the economic success of I&PI countries.

Perhaps the best estimates of national levels of cognitive abilities that are currently available were developed by the German psychologist Heiner Rindermann (2007, 2008; Rindermann & Thompson, 2011). Rindermann combined GCA test scores and educational achievement measures into a single national estimate of cognitive abilities. The global distribution of Rindermann’s scores was similar to the distributions reported by Lynn and Vanhanen. In general, high scores were obtained in Europe and in countries whose culture is largely derivative from Europe, such as the United States and Australia, and in the industrial Asian countries. The lowest scores were obtained in sub-Saharan Africa. Nations in Latin America, nonindustrialized Asian countries, and nations in the Middle East and North Africa had intermediate scores.

Rindermann correlated his scores with a number of other indices of national well-being, ranging from GDP/c to the rate of homicides. Some of the findings are shown in Table 4. In general, properties of society that are considered desirable in I&PI societies, such as a high GDP/c, efficient government bureaucracy, and effective policing, were associated with high cognitive ability scores. For instance, indices of the rule of law and of GDP/c had a correlation of slightly above .6 with the cognitive index. Properties that are considered undesirable in the I&PI society, such as a high incidence of HIV/AIDS or a high rate of homicides, were associated with low cognitive scores.

The negative correlation between cognitive scores and fertility, −.7, was the highest absolute association in Rindermann’s (2006) data set. This is a particularly interesting statistic, for it illustrates a marked difference in attitude between developed and developing countries. Among countries that are members of the Organisation for Economic Co-operation and Development (OECD; which contains most I&PI countries and a few that are close to I&PI status, such as Chile, Mexico, and Turkey), the mean desired number of children at the start of the 21st century was 2.25, only slightly above the replacement rate for the population. In Nigeria, the mean desired family size was 6.7 children (Westoff, 2010). The difference is important, because high levels of fertility, and the resulting high levels of family size, present a society with economic and social pressures that work against the development of effective modern education systems. In addition, large families place limits on women’s socioeconomic roles. Both good schooling and substantial gender equity are required if a society is to participate fully in the economics of the I&PI world.

Rindermann’s (2006) analyses were based on data collected from the 1970s until 2000. I updated his results by calculating the correlation between 2009 GDP/c and the 2009 PISA results. Over all data, the correlation was only .32. However, three small countries, the oil-rich emirate of Qatar and the duchies of Lichtenstein and Luxembourg, both of which are major financial centers, had unusually high GDP/c relative to their PISA scores. Removing these three atypical states from the analysis raised the correlation to .65, which was surprisingly consistent with the results of earlier studies.

Correlation does not mean causation is a mantra of statistical analysis. Rindermann (2006, Figure 5) was able to construct a longitudinal analysis based on 17 countries for which he had data in both the 1970s and the early 2000s. The relative standing of countries was fairly stable, on both the cognitive variables and GDP/c. The standardized regression measures were in the middle .7 range. In general, the rich stayed rich, and the smart stayed smart. Conversely, there were interactive effects: GCA scores in the 1970s predicted GDP/c in 2000...

<table>
<thead>
<tr>
<th>Desirable</th>
<th>Correlation</th>
<th>Undesirable</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule of law</td>
<td>.64</td>
<td>Fertility rate</td>
<td>−.73</td>
</tr>
<tr>
<td>Quality of bureaucracy</td>
<td>.64</td>
<td>Gini economic inequality index</td>
<td>−.51</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>.63</td>
<td>HIV infection rate</td>
<td>−.48</td>
</tr>
<tr>
<td>Economic freedom</td>
<td>.52</td>
<td>Government spending as percentage of economy</td>
<td>−.47</td>
</tr>
<tr>
<td>Economic growth</td>
<td>.44</td>
<td>Homicide rate</td>
<td>−.23</td>
</tr>
<tr>
<td>Rate of solved homicides</td>
<td>.32</td>
<td>War: frequency and impact</td>
<td>−.22</td>
</tr>
</tbody>
</table>

Note. Data are from Rindermann (2008, Figure 2). GDP = gross domestic product.
intelligence in Figure 1A. A nation, or any other population that acts as a unit, can exert control over its physical and social environment in more powerful ways than an individual, working alone, can. The following examples illustrate feedback processes in which social policies have influenced the physical and social environments and even the gene pool itself:

1. **Fetal alcohol syndrome and related effects**: Fetal alcohol syndrome is a serious cognitive disorder exhibited by children of women who abused alcohol during their pregnancy. Milder forms of the disease are believed to be associated with alcohol use below the level normally considered as abuse. Many nations now conduct extensive public health campaigns to discourage pregnant women from consuming alcohol. As a physical variable is involved, the effect of this campaign is to modify information processing capacities.

2. **The Flynn effect**: During the 20th century, cognitive test scores increased sharply in industrially developed nations (Flynn, 2007). Although the causes for this are not clear, surely one contributing factor was the increased availability of, and requirement for, education through the high school level. As in the case of gender differences, increased availability of education throughout the population will have had an influence upon the development of the population potential for GCA.

(c) **Assortative mating**: Assortative mating is a nonrandom tendency toward mating between people with similar phenotypes—for instance, tall men marrying tall women. In the industrially developed countries, there is a correlation of approximately .3 between the cognitive test scores of spouses, which indicates a modest amount of assortative mating for GCA (van Leeuwen, van den Berg, & Boomsma, 2008). It is important to note that although the concept of assortative mating arose in genetics, the phenomenon can be produced by either biological or purely social processes. For instance, in the United States, adult social contacts are strongly related to education; people with similar educational levels socialize together and hence are likely to have children together. Because the educational system in the United States acts as an imperfect “sorting machine” for general cognitive abilities, assortative mating is, to some degree, virtually guaranteed by the organization of society (Murray, 2012). As children receive both their genetic inheritance and, in the vast majority of cases, their social inheritance from their parents, assortative mating will operate to increase the variance in GCA, across generations, through both social and genetic mechanisms.

The development of national GCA relies both on the incorporation of cognitive artifacts from the nation’s own history and from the societies with which it is in contact. The “Arab Spring” revolutions in 2011 provided a striking example. The rebels who brought down the reactionary Egyptian, Libyan, and Tunisian regimes relied heavily on cell phones to coordinate their activities.

And where did the cell phone come from? Cell phones were developed in Europe and North America. One of the basic concepts behind the technology, high-speed frequency switching, was first proposed in a U.S. patent filed in the 1940s by two Austrian refugees from the Nazi expansion. One of the Austrians was the motion picture actress Hedy Lamarr (Rhodes, 2011). It pays for a society to be alert to new cognitive artifacts, wherever they may be found.

### Genetic differences between national populations

**Direct tests of the hypothesis that there are genetic differences between nations**. Gene frequencies vary across populations. Sub-Saharan Africans do not look like Scandinavians, and many of their differences in appearance, such as eye color, are dictated by genetics. Other differences, such as language spoken, have nothing to do with genetics. Still other differences, such as weight, are partially influenced by genetics. Determining the extent to which the genetic potential for GCA differs across national contrasts is a challenging task, for all we can measure is expressed GCA, and the features of a society that convert genetic potential into expressed GCA vary widely across nations.

It is well established that across individuals, variations on cognitive test scores are correlated with variation in genotypes. This point was originally made in the 1960s, through studies of correlations between individuals of different degrees of relation. This finding has since been replicated many times, using a variety of statistical models (Hunt, 2011, Chapter 8). Many genes are involved, with no one gene accounting for a large amount of the variance (Davies et al., 2011; Johnson, Penke, & Spinath, 2011). As of 2012, the genes involved are unknown. Given the genomes of two individuals, it would be possible to make an estimate of how far apart their intelligence test scores were likely to be, but it would not be possible to predict which of the two had the higher scores or for that matter the mean score of the pair.

By comparing gene frequencies across populations, it is possible to estimate the degree to which two population groups are genetically related. A number of the genes that vary across racial–ethnic groups have been identified. However, because the genes involved in establishing the potential for cognitive ability are not known, there is no way of predicting which of two populations has the higher genetic potential. This is not necessarily a permanent situation. Future developments in molecular and behavior genetics may identify these genes, and
at that time it will be possible to make a scientifically justifi-
able statement about international differences in the potential
for intelligence. As of 2012, no such statement can be made.
This does not mean that no such differences exist. It means
that the extent of a genetic contribution to international differ-
ences in GCA is unknown at present.

The analogical argument for genetic causes of national
differences. Arguments for large genetic influences on inter-
national differences in intelligence are often based on an anal-
ogy to intragroup evidence for the heritability of intelligence
test scores observed in U.S. and European studies, most of
which have contained only White participants. Within these
populations, the heritability coefficient for intelligence is
somewhere in the .5–.8 range. (See Hunt, 2011, Chapter 8, for
a discussion of the many studies of related individuals that
lead to this conclusion and Davies et al., 2011, for a study that
reaches the same conclusion on the basis of correlations of
-genetic makeup and intelligence test profiles in unrelated indi-
viduals.) The analogical argument is that because intelligence
has a genetic basis within White U.S. and European popula-
tions, and because international populations differ genetically
and in general cognitive ability, is it not reasonable to believe
that international differences in cognitive ability are due to dif-
fences in genetic potential? The same argument has been
used to explain racial differences in test scores across racial
and ethnic lines within the United States (Lynn & Vanhanen,

The argument has both a logical and an empirical weak-
ness. Consider several groups and two measured variables, A
and B, on individuals within each group. A correlation between
A and B within groups does not imply that the correlation will
be found between groups. It may be or it may not be. There-
fore the analogy is logically weak. The analogy’s empirical
weakness is a consequence of the definition of the heritability
coefficient. Heritability coefficients are measures of the rela-
tive influence of genetic and environmental differences on a
trait. The value of a heritability coefficient is determined by
three things: the proportion of individuals who express a phe-
notypical trait given a constant genetic makeup (technically,
penetrance), the extent of relevant genetic variation within the
population under investigation, and the extent of relevant
environmental variation. Eye color has a penetrance close to 1.
Heritability coefficients for such traits are close to 1, every-
where. Other traits, certainly including cognitive ability, have
much lower penetrances, and therefore heritability coefficients
can vary across populations. Height is an excellent example.
In Finland, Australia, and the European-derived populations
within the United States, the heritability coefficient for height
is .80. Almost everyone in these populations has access to
adequate infant and early childhood nutrition programs, which
are a major environmental influence on adult height. In Asia
and Africa, where the genetic variation is different and, most
important, the adequacy of nutrition varies widely, heritability
coefficients fall to .65 (Lai, 2006). In an extreme case, North
and South Koreans are genetically close populations, having
intermingled freely for centuries prior to their separation into
two states in 1945. Today South Korea is a prosperous mem-
ber of the OECD, while North Korea has suffered major fam-
inges. By their late teenage years, South Korean men are 16 cm
taller and South Korean women are 13 cm taller than their
North Korean counterparts (Pak, 2010; Schwekendiek, 2009).

Cognitive ability is certainly more like height than eye
color. The vast majority of the studies establishing the .50–.80
estimate for the heritability of intelligence were conducted in
middle-class populations in I&PI nations, with widespread
health and nutrition programs, relatively stable home environ-
ments, and extensive educational programs. Studies of low
-socioeconomic status (SES) groups in the United States find
much lower heritability coefficients (Turkheimer, Haley,
Waldron, D’Onofrio, & Gottesman, 2003). This may be due to
greater relevant environmental variability in these groups than
in the middle-class participants in the earlier studies.

The environmental differences between nations appear to be
much larger than within-nation differences, especially when the
contrast is between exceptionally low-scoring nations such as
those in sub-Saharan Africa, and high scoring nations, such as
those in Scandinavia. Simply generalizing the high heritability
coefficients observed within the I&PI nations to estimate the
role of genetics in variation across nations is not war-
anted. The conclusion is not logically implied, and there are
reasons to suspect the empirical basis. However, the analogy
cannot be denied either.

The argument from evolutionary necessity. A second indi-
 rect argument for international differences in the genetic
potential for intelligence is based on evolutionary psychology.
The argument is that following the “out of Africa” exodus, dif-
ferent groups ancestral to modern populations faced different
environmental challenges as they moved through Eurasia,
Oceania, and eventually the Americas. According to these
hypotheses, within those populations that moved through cogni-
tively challenging environments, there was a reproductive
advantage to being intelligent. This could have arisen either
from selection pressures that increased the frequency of alleles
favorable to the development of intelligence in an existing
gene pool or by a mutation that increased the survivability of a
group migrating through a hostile environment. Analogies are
sometimes drawn to the fact that mutations producing light
skin color did occur, separately, in prehistoric North Asian and
European populations. The mutations produced more effective
utilization of sunlight to produce Vitamin D at the expense of
greater sensitivity to sunburn. If such a favorable mutation
could occur for the skin, why not assume that a similar muta-
tion occurred for cognition?

Lynn (2006) has proposed that the European and North
Asian ancestors of present-day populations faced harsh, cold
climatological differences that exerted selection pressures for
cognitive abilities. Such challenges, and the resulting selection
pressures, were, according to Lynn, not as severe in the warmer
savannas of East Africa. Kanazawa (2004) has contended that selection pressures for GCA occurred as the ancestral populations moved through different environments. For example, the populations ancestral to modern North Asians had to move from the savannas of Ethiopia, as they may have been 80,000 years ago, through the arid, hot deserts of Arabia and the Middle East, and then across mountain ranges to reach the colder, better watered areas of present-day northern China.

Both these arguments are plausible, when considered alone. A more critical examination, though, reveals several weaknesses. To me, the strongest is that the evolutionary argument is an argument for the presence of differences in the frequency of genes related to intelligence, and these differences have yet to be shown to exist. Lacking such direct evidence, both Kanazawa (2004) and Lynn (2006) draw inferences based on present-day distributions of test scores. Wicherts, Borsboom, and Dolan (2010) pointed out two problems in using these data to draw conclusions about evolutionary processes. The analyses presented by Kanazawa and Lynn are heavily influenced by the disparity between the very low sub-Saharan African scores reported by Lynn and Vanhanen (2002, 2006). These are spuriously low. Even if they were not, the two hypotheses implicitly assume that the distribution of the genes for cognitive potential was approximately the same in the ancestral populations, some 6,000 to 10,000 years ago, as it is today. There is no evidence that this is so.

The support for an evolutionary explanation of present-day differences in national measures of GCA is weak at best. Definitive evidence for or against this hypothesis would require comparative data on group cognitive behavior in prehistoric times. It is unlikely that such evidence will ever be obtained.

A summary evaluation of the genetic hypothesis. The genetic hypothesis states that international differences in the genetic potential for intelligence exist and that they are driven by differences in the relative frequencies of the genes for intelligence in various international populations. Because the genes that establish a potential for intelligence are not known, there is no way to test the hypothesis directly. Indirect arguments for the hypothesis have been made, by analogy to the results of studies of the genetics of intelligence within national populations and by proposing evolutionary theories of intelligence. Both these arguments contain substantial weaknesses.

If, at some time in the future, the genetic basis for intelligence is determined, it will be possible to test the genetic hypothesis for group differences directly. Such tests could be made both internationally and across population groups within a nation. Until then, the question “Is there a genetic basis for international differences in intelligence?” has a simple answer: We do not know.

The physical environment

Differences in the physical environment fall into three broad categories: nutrition, health, and environmental pollutants. The effects of problems in these areas on the development of individual intelligence are well known (Hunt, 2011). What are the international effects?

Nutrition. There are huge international differences in nutrition and health. On a given day, a randomly chosen sub-Saharan African is more than 10 times as likely to have two or fewer meals per day as is a randomly chosen European. Table 5 presents some illustrative data for five countries. Brazil is something of a special case, for it is often pointed to as a country whose socioeconomic trajectory is taking it into the I&PI group.

Table 5. Rates of Stunting and Disability-Adjusted Life Years in Five Selected Countries

<table>
<thead>
<tr>
<th>Country</th>
<th>GDP per capita corrected for local purchasing power (rank of 193)</th>
<th>Rate of stunting (%) in children under 5 years of age</th>
<th>Disability-adjusted life years (incidence per 1,000 people)</th>
<th>GCA estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>$39,171 (13)</td>
<td>negligible</td>
<td>10.32</td>
<td>99</td>
</tr>
<tr>
<td>Brazil</td>
<td>$11,273 (76)</td>
<td>7</td>
<td>20.11</td>
<td>87</td>
</tr>
<tr>
<td>Egypt</td>
<td>$6,417 (104)</td>
<td>29</td>
<td>20.26</td>
<td>84</td>
</tr>
<tr>
<td>India</td>
<td>$3,408 (129)</td>
<td>48</td>
<td>27.82</td>
<td>79</td>
</tr>
<tr>
<td>Nigeria</td>
<td>$2,437 (141)</td>
<td>43</td>
<td>48.58</td>
<td>69 (77)</td>
</tr>
</tbody>
</table>

Note. GCA = general cognitive ability.

1 Data are from UNICEF Statistics (http://www.unicef.org/info/bycountry).
2 Data are from the World Health Organization, 2011 (http://gamapserver.who.int/gho/interactive_charts/mbd/as_daly_rates/atlas.html).
3 Estimates are from Lynn and Vanhanen (2006).
4 The parenthetical value is Wicherts, Dolan, Carlson, and Van der Maass's (2010) correction, which is probably a more accurate estimate than the Lynn and Vanhanen estimate.
Stunting is caused by very poor nutrition in early childhood. Not surprisingly, stunting is also associated with reduced cognitive development (e.g., Crookston et al., 2011). Stunting is an unusual condition in the first world. In Canada, the rate is negligible. In Brazil, where virtually all socioeconomic indicators have improved in the last 20 years, there is still some stunting. Egypt’s rate is 4 times that of Brazil. More than 40% of the children in India and Nigeria are stunted, according to international standards.

These examples are illustrative of general trends. The Lynn and Vanhanen (2006) estimates of national GCA are sharply divided between the high-scoring developed nations and low-scoring developing nations (Hunt & Wittmann, 2008). Stunting rates are less than 5% in the countries of the I&PI world and 23% in countries outside of it. The difference in stunting rates provides striking evidence of the differences in level of childhood malnutrition worldwide.

Childhood malnutrition influences cognition in three different ways. In the case of extreme and prolonged malnutrition, there are direct effects on the central nervous system. There are also substantial interactions with other influences on cognition. Malnutrition increases susceptibility to infectious diseases, which themselves may damage the nervous system. Malnutrition also interacts with education, for malnourished children have difficulty concentrating in educational settings. Looked at more positively, good nutrition provides the attentional capacity needed to benefit from education. For instance, studies in Latin America have found that nutritional supplementation programs have their greatest impact when educational interventions are also provided, and vice versa (Pollitt et al., 1993).

Persistent hunger, far short of the level one would describe as malnourishment, also has an effect on the social environment outside the school. When the family is hungry, it makes sense to have the children work. Work may vary from tending goats on the savannah to selling candy, cigarettes, and the like on the streets of overcrowded third-world cities. Children certainly learn some things while they are working, but they are not learning how to use many of the cognitive artifacts required in modern society.

**Health.** The second environmental variable shown in Table 5 is the WHO disability-adjusted life years (DALY) index. This is a measure of the loss of useful years due to health conditions, including infectious diseases, chronic diseases such as diabetes, and behavioral practices (e.g., tobacco use). The index reflects incidence, severity, and age of onset of the various conditions. Therefore, the DALY is a measure of health in the general population, not just in children. Table 5 presents the DALY index for the five countries used to illustrate differences in nutrition. Again the discrepancies are striking.

There is a strong, negative relation between national measures of GCA and the DALY index. Eppig, Fincher, and Thornhill (2010) report a correlation of –.76 between the logarithm of the DALY index and Lynn and Vanhanen’s (2006) national GCA estimates, after adjusting for a probable bias downward in the sub-Saharan countries. The DALY can be broken down into separate indices reflecting the contributions of major infectious diseases and the contributions reflecting nutritional deficiencies. The partial correlation between the IQ estimates and the DALY infectious disease estimates, statistically equating the effect of nutrition on test scores, was –.56. This indicates that infectious diseases exert a negative influence on national intelligence levels, beyond the effects of nutrition alone.

Although nutrition and health do act on the national potential for GCA, a simple “health causes a potential for intelligence” model is naive. There are strong feedbacks. The development of national public health and nutrition systems itself implies the existence of appropriately educated professionals and subprofessionals. Health professionals are of little use unless they are backed by national priorities and the commitment of wealth sufficient to create the nutrition and health systems. This implies an interested, sophisticated, and stable government. The population itself has to have sufficient understanding to be convinced to adopt practices that support better health, especially if the new practices interfere with less effective traditional practices. Take a historic example. In imperial Rome, water quality standards were based on qualities that can be directly sensed, such as taste, clarity, and smell. A proposal that otherwise healthy appearing water not be used because it contained invisible, highly dangerous microorganisms would probably have been regarded as contrary to the public interest. The germ theory of disease is a modern, useful cognitive artifact.

**Environmental pollutants.** Environmental pollutants, and especially heavy metals, affect cognition adversely. In the 17th and 18th centuries, the expression “mad as a hatter” reflected the use of mercury in manufacturing hats and the resulting poisoning that prolonged exposure induced. Today, lead is one of the most potent of present-day atmospheric pollutants (Hubbs-Tait, Nation, Krebs, & Bellinger, 2005), in spite of a two-millennia history of attempts to control it (Hunt, 2011, Chapter 9). The modern history and situation with respect to lead presents a case study in the interplay between politics, economics, and health.

Tetraethyl lead was introduced as an anti-knock agent in gasoline in the 1920s. It continued to be used in I&PI countries until the 1970s, in spite of documentation of its effects on cognition. Lead-added gasoline is still used in many developing nations (Tong, von Schimding, & Prapamontol, 2000). Non-atmospheric sources of lead poisoning, such as lead-based paint or lead used in industrial production, are substantially more common outside than inside I&PI nations. However, I&PI nations are far from immune, because of product importation from nonindustrial to I&PI countries. In 2007, the Mattel toy manufacturing company recalled 1.5 million Chinese-made toys because of excess lead in paint. This cost the company $30 million dollars, approximately half of its quarter’s operating profit. At a more prosaic, retail level, before American tourists to Latin America bring home the elegant,
inexpensive pottery they can find in the market, they would be well advised to enquire about the method of glazing used.

A case study by Karla Guerrero Leiva (2009) shows how serious the problem is. She obtained blood samples from children in a primary school in a poor district of Callao, a port city adjacent to Lima, Peru. The school was virtually on the same block as an industrial lead deposit. The mean concentration of lead in the children’s blood was 40.7 micrograms per deciliter. What does this mean?

For reference, the WHO recommended maximum level is 10 micrograms per deciliter, although decrements in intelligence have been reported for young children with blood levels lower than this (Canfield et al., 2003). The mean observed level in Mexican American children in the United States is 4.4 micrograms per deciliter. Less than 3% of this population has blood levels above 7 micrograms per deciliter (Schwemberger et al., 2005). The distribution of levels in the Peruvian study is shown in Table 6, along with some selected symptoms associated with each level of lead-serum concentration. Every one of the 66 children in Guerrero Leiva’s study would have been considered a candidate for medical intervention if they had lived in the United States.

The Peruvian case is not an isolated example. Excessive exposure to heavy metal pollutants is considered a substantial risk factor throughout the developing countries of the world (Walker et al., 2007).

As was the case for health and nutrition, the interaction between pollution and intelligence, at the population level, is not a simple one. In general, the effects of pollution on the atmosphere are more easily sensed than the effects of microorganisms in the air and water. The problems arise more from complex tradeoffs between economic needs and the costs of either altering or eliminating an economically viable but environmentally hostile activity. Solving such problems requires an educated, cognitively skilled population or, at least, a cognitively skilled cadre of individuals who have the ability to understand the problem and also have the motivation to solve it.

**The social environment**

According to Diane Halpern (Halpern et al., 2007), psychology has a law that is as true as the law of gravity: *You learn to do what you practice doing.* If you practice thinking, you will learn to think. If you do not, you will not. The opportunity to practice thinking begins with the home, progresses through school, and is then supported, or suppressed, by exposure to and attitudes toward innovation. The development of national GCA proceeds on this principle. Does the society provide an environment that encourages thinking and exploration of ideas?

**The home environment.** Within countries, ratings of home environments have been shown to be substantial predictors of scores on tests of cognitive skills. The evidence for this is based on studies ranging from Ohio to the Philippines (Hunt, 2011, Chapter 11). It is difficult to make cross-cultural comparisons of home environments, because cultures differ so greatly in home structure and practices. However, we can say a few things.

- Family size is important. As pointed out earlier, at the international level, fertility rates (and hence large families) are negatively correlated with measures of cognitive skill. A similar relationship has been observed at the individual level in many studies in developed countries: Women with low test scores tend to have more children than do women with high test scores. Because parental and child scores are positively correlated, it follows that children from large families will tend to have low scores. This could be either because the children in large families receive reduced genetic potentials from their parents or because the parents, due either to low cognitive skills or to the sheer burden of dealing with a large family, provide less than ideal circumstances for child rearing.

- Proponents of genetic explanations for international differences in intelligence emphasize the genetic possibility. It has even been claimed that because of the higher fertility rates of countries with low test scores, the worldwide genetic potential declined by 0.86 IQ points in the years 1950–2000 and will decline by a further 1.28 points between 2000 and 2050 (Lynn & Harvey, 2008). It is a bit hard to take this argument seriously, for we cannot measure genetic potential directly, observed test scores rose during the 20th century, and the authors did not consider the changes in fertility rates that are occurring worldwide. Nevertheless, it is true that to the extent that there are international differences in the genetic potential for intelligence, which could be the case, although it is far

---

**Table 6.** Lead Levels Observed in Children in a Primary School Near an Industrial Lead Deposit in Callao, Peru (Guerrero Leiva, 2009)

<table>
<thead>
<tr>
<th>Lead level in blood (micrograms per deciliter)</th>
<th>Psychologically relevant symptoms expected (Tong, von Schimding, &amp; Prapamontol, 2000)</th>
<th>Number of children at indicated level</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–10</td>
<td>Not considered poisoned</td>
<td>0</td>
</tr>
<tr>
<td>10–20</td>
<td>Slowed growth of CNS</td>
<td>5</td>
</tr>
<tr>
<td>20–40</td>
<td>Lowered nerve velocity</td>
<td>36</td>
</tr>
<tr>
<td>40–70</td>
<td>Serious CNS symptoms, elevated coproporphyrin in urine</td>
<td>25</td>
</tr>
</tbody>
</table>

*Note. CNS = central nervous system.*
from proven, differential fertility rates would accentuate the differences between nations over time.

Family size alone matters, apart from genetics. Strong evidence for this proposition has been obtained in Norwegian studies of birth order effects (Bjerkedal, Kristensen, Skjeret, & Brevik, 2007; Kristensen & Bjerkedal, 2007). In Norway, all men are required to register for military service at age 18. They are given a cognitive test at that time. The resulting large database, stretching over a number of years, makes it possible to test several hypotheses concerning both birth order and family effects. One result is particularly compelling. In the overall database, the mean IQ equivalent score of first-born men was 103, and the mean for second-born men was 100. However, the mean score of second-born men from families where the first-born died in childhood was 103. Similar increases in scores were observed in third-born children: an IQ of 99 in a family with no deaths of older siblings, 100 with one death, and 103 with two deaths.

Why should family size be so important? In the poorer nations in the world, and in low SES groups in the richer nations, the need to support many children forces adults to focus on the family’s economic well-being. This requirement competes with time for adult-child interactions that facilitate the development of the child’s mind. The effect is exacerbated when children are raised in understaffed orphanages or in families that are headed by an older child. Both these conditions are much more frequent in the developing world than the industrial world. The absence of adult heads of family is an acute problem for counties that have high incidences of HIV/AIDS in the adult heterosexual population, for this disease strikes adults at the time in their lives at which they would normally be caregivers for their children.

Families are also important as motivators for educational achievement. As children mature, the family has to convince the student that educational effort is worthwhile. This is especially true for abstract but crucial cognitive artifacts. How often does an eighth grader see someone using trigonometry? Nevertheless, trigonometry homework is important, and the family environment must encourage it. For, as Halpern et al. (2007) have explained, if you don’t practice, you won’t learn. If the family structure is not oriented toward encouraging students to work at learning throughout their career as students, the amount learned will be reduced.

**Formal education.** A great many things can be learned outside of school in every society. Nevertheless, formal education is the way that students acquire skills in using the cognitive artifacts of the I&PI society.

Education is a cumulative process. The amount a child learns in the first grade determines the extent that he or she is ready to participate in the second grade, the amount learned in the second grade influences the amount learned in the third grade, and on and on. U.S. studies have shown that the extent to which a child has specific skills, such as prereading skills, and general habits facilitating the control of attention prior to entering primary school will influence the rate at which he or she acquires information in later grades, extending through high school (Phillips, Crouse, & Ralph, 1998). Longitudinal studies such as this are fully consistent with PISA reports of correlations between ratings of the home educational environment and the academic achievement of 15-year-olds. Once again on the basis of U.S. data, it is now clear that handicaps due to a lack of family support for the acquisition of school-relevant cognitive skills can be partially ameliorated by preschool education programs. In spite of early pessimism (e.g., Jensen, 1969), these programs do work and are cost-effective, although recovering the cost of the investment takes decades (Barnett, 2011; Reynolds, Temple, Ou, Artega, & White, 2011). Similar early intervention programs are available within other I&PI nations. They are few and far between in the developing countries of Africa, Asia, and Latin America.

The sheer amount of formal education matters. Analyses within the developed world suggest that each year of education is associated with a rise in intelligence test scores equivalent to 2.7 points on the IQ scale (Winship & Korenman, 1997). If this is so, we should expect differences in educational requirements to be reflected in differences in a country’s socioeconomic success. They are. Rindermann (2008) developed a structural equation model in which, among other things, he determined the standardized regression coefficient for the number of years of education in a country in the 1970s and the country’s GDP/c in 2000. The coefficient was .40, which is of considerable economic importance.

Expanding on Rindermann’s (2008) idea, the key statistic is the expected number of school years that a child can anticipate upon entering first grade. This is the educational analogue of the familiar life expectancy measure, which is the (estimated) number of years between birth and death. In the sub-Saharan nations of Africa, the expected number of school years is 9.5 years, with a range from 5 (Eritrea) to 13.5 (the island nation of the Seychelles). By contrast, the five developed nations of Canada, France, Germany, the United Kingdom, and the United States average 16.5 years with essentially no variation.12

A thought experiment shows how important these discrepancies might be. Jetje Wicherts and his colleagues (Wicherts, Dolan, Carlson, & Van Der Maas, 2010a; Wicherts, Dolan, & Van Der Maas, 2009) have estimated that sub-Saharan African nations have a mean intelligence test score in the 80–82 range.13 Suppose we set the five developed nations’ scores at 100. There is, on the average, a difference of 7 years in the expected number of school years between the African and the developed nations. Applying Winship and Korenmann’s (1997) 2.7 IQ points gain per year estimate, the difference in school years between the sub-Saharan and developed nations would account for an 18.9 IQ point difference—enough to wipe out the difference in reported national intelligence scores.

One should not take this exercise literally, for it assumes that the quality of the added schooling and the social support for students are equivalent in the developing nations of Africa and in postindustrial Northern Europe and North America.
Rindermann and Ceci (2009) addressed this issue, at the international level, by computing causal models relating the characteristics of national school systems to performance on international assessment programs, such as PISA. They found that the following characteristics lead to good performance, at the national level:

1. By far the most important predictor of the success of current students was the general educational level of the country, as indexed by such things as adult literacy and various other indices of adult academic achievement. This is consistent with Rindermann’s (2006) earlier finding that “the smart get smarter” over time. However, it presents a policy maker with something of a conundrum. The way to have a good educational system is to embed it in a country that already values and achieves educational success. Such values imply a high rate of participation among the current school-age population, something that Rindermann and Ceci (2009) found to be a predictor of high national achievement.

2. The second most important predictor is early entrance into school, particularly kindergarten. This appears to be at least partly because early entry facilitates the development of school-relevant habits, such as focusing attention.

3. Countries that make liberal use of formal assessment procedures and that track students into ability groups fairly early in the students’ careers perform better on international assessments of progress than countries that do not.

Rindermann and Ceci (2009) argued that national educational policies could exert a major influence on all of these variables. Let us look at some of the problems that policymakers face.

In the I&PI countries, access is seldom not an issue; virtually all students can attend school up until they are 15 or 16. In general, the majority of school fees are paid by the state, not the family. Dropout problems exist, but these are largely due to pressures on the individual student and mainly, though not exclusively, affect lower SES groups. The default alternative is that a child will go to school.

Outside the I&PI world, there are structural disincentives to schooling. Poorer countries shift the cost of primary and secondary education to students and their families much more than the postindustrial countries do. For instance, on the Tanzanian mainland, the gross national income per capita in 2011 was $1500, while high school fees, including fees and uniforms, cost about $300. In other cases, children may be taken out of school because the family needs their financial contribution, which may vary from working as a herder on the African veldt to selling trinkets on the streets of increasingly crowded third-world cities. These are learning experiences, but they are not equivalent to formal schooling.

Rindermann and Ceci’s (2009) point about the importance of kindergarten assumes that the kindergarten experience actually does prepare the child for school. The quality of kindergarten and primary school education is more sensitive to student–teacher ratios than is secondary education, and the ability to maintain an appropriate student–teacher ratio is dependent on the availability of trained teachers. The problem is exacerbated if teachers’ working conditions are poorest in those sections of a country most in need of primary education. This is often the case, even in I&PI nations. Some countries attempt to solve the problem by requiring newly credentialed teachers to spend a certain number of years in an affected area, such as a poor area of a city or in impoverished rural areas. To the extent that these teachers leave their assigned positions as soon as they can, the most vulnerable children are being educated by the least experienced teachers.

The use of frequent examinations and tracking raises some interesting cultural issues. Examinations do three things: evaluate individual students’ learning, provide feedback to teachers and administrators about how the system is doing (rather like quality control studies in industry), and serve as management tools, to indicate to teachers what students are expected to learn and, by inference, what teachers are expected to teach.

Public education systems that practice tracking typically use examinations to group students by their current educational level and then offer relatively uniform programs to a relatively uniform group of students. The most common form of tracking in I&PI nations is to group students into “academic” and “vocational” tracks. The provision of “special education” programs for students having major academic difficulties and (usually much smaller) advanced programs for high-achieving students are also forms of tracking. The hope is that if teachers in each track have to deal with relatively uniform students with similar educational goals, overall academic achievement will be higher than if there is a high degree of variation in student goals and achievements within each class. Rindermann and Ceci’s (2009) analysis suggests that this hope can be realized.

Outside I&PI countries, early tracking based on academic performance is less prevalent. However, socioeconomic tracking occurs to the extent that parents are required to pay for the education or, in some areas, by the simple availability of schools. Tracking at the high school level is widespread, because poorer countries do not have enough secondary school facilities for the entire population. As a result, assessment programs that focus only on students enrolled in the schools, such as PISA, overestimate the level of I&PI relevant cognitive skills in the country. This does not mean that the people not assessed are stupid or that they do not have cognitive skills. It does mean that they have not been trained to exercise the cognitive skills and use the cognitive artifacts that are important in industrial societies.

The role of examinations in providing feedback and guidance to educators is too often overlooked. The content of a formal examination can be thought of as a message telling...
teachers and principals what they should be teaching. For instance, the U.S. No Child Left Behind (NCLB) program focuses on language usage, mathematics, and science. The international PISA program has a similar structure. In order to prepare students for such programs, teachers should ensure students’ knowledge of language arts and mathematics, and because of the breadth of the science category, students need to have a broad familiarity with topics in most of the sciences but need not have a deep understanding in any one of them. Student knowledge of humanities and arts is irrelevant to performance on the NCLB examinations. By contrast, the far more extensive (and more demanding of student time) British General Certificate of Secondary Education examination program tests the same topics as the NCLB and in addition offers students choices of examinations in the individual sciences (thus evaluating deeper knowledge in a narrower area of science) and various topics in the humanities and arts field. Therefore, the U.S. examination system encourages teachers to focus their instruction narrowly. The British system encourages broader instruction.

Other I&PI countries generally fall in between the British and American systems. Most of the developing countries simply do not have the luxury of a choice, with the exception of offering courses that focus on their own history and culture. Therefore, education to impart the cognitive skills required by Western society may be limited to mathematics and fairly rudimentary science.

**Education at the top.** At the national level, the development of human resources is not just a matter of ensuring minimal education to all, as evidenced by high literacy rates and low dropout rates. Performance at the top counts a great deal. National measures of high-level academic achievement, such as the frequency of college graduates or the test scores of the top 5% of the student population, better predict economic success than do more general indicators, such as literacy rate (Rindermann & Thompson, 2011). Performance in mathematics is particularly important because of mathematics’ centrality in the science, technology, education, and mathematics (STEM) fields. In order to meet national needs, the best students have to be encouraged to do more than get by—they need to be encouraged to become experts. Achieving high levels of performance requires a lot of hard work. Cultural practices influence both the opportunity students have to put in the time needed to acquire high-level skills and the encouragement students are given to do so.

Figure 2 shows the 2009 PISA mathematics scores for 15-year-old boys and girls in the 10 (out of 65) highest scoring countries. (The United States, which is not shown, ranked 31st overall with scores of 477 for girls and 497 for boys.) There was considerable variation between countries, even within the top 10. Boys outperformed girls in every one of the top 10 countries (and in the 65 countries generally). However, the gender differences were far smaller than the international differences. The 15-year-old girls in Singapore outscored 15-year-old boys in the other 64 countries.

The top 10 countries included four industrialized Asian countries: South Korea, Singapore, Taiwan (China–Taipei), and three industrialized autonomous regions of the People’s Republic of China. Explanations offered for Asian prominence in mathematics have run the gamut from presumed genetic differences to Confucian respect for scholarship. The PISA data, and the observations of many who have visited Asian communities, are consistent with a simple behavioral explanation: Asians and students in other high-scoring nations simply spend more time practicing mathematics than do students in other countries. In the United States, the school year is from 34 to 36 weeks long, depending on the state. European school years are generally 2 weeks longer. In the Asian countries shown in Figure 2, school years are 40 to 43 weeks long. In addition, students in the industrial Asian countries are considerably more likely to receive additional mathematical instruction than are students in predominantly non-Asian countries. This is shown in Figure 3A.


**The importance of attitude.** In recent years, a great deal of research has documented the importance of attitude and motivation as determinants of cognitive accomplishment (Duckworth, Quinn, Lynam, Loeber, & Stouthamer-Loeber, 2011). This is not inconsistent with an emphasis on practice, for by and large a person will practice those activities that he or she regards as culturally appropriate and potentially rewarding. Two further aspects of international differences in students’ accomplishments in mathematics are relevant.

The gender difference in mathematics shown in Figure 2 has been found in previous PISA examinations. The 2003 results are of special interest, for these assessments concentrated on mathematics. Various measures of gender inequality had correlations in the .5 region with the size of the gender gap. The more opportunity women had to use their education, the better their educational achievement (Else-Quest, Hyde, & Linn, 2010; Guiso, Monte, Sapienza, & Zingales, 2008). This...
toward educational achievement. The upper panel of Figure 3 shows that students in industrialized Asian countries received more out-of-class instruction in mathematics than did students in non-Asian I&PI countries. The lower panel shows a qualitative difference in the way that the extra instruction was used. In European and North American countries, the students who received extracurricular instruction did worse on the examination than the students who did not. It is not likely that this was because Western extracurricular instruction was so bad that it did harm. What is far more likely is that the Western countries used extracurricular instruction to support weaker students, in order to avoid failure. This is consistent with the philosophy suggested by the title of the major U.S. educational effort of 2001–2011, “No Child Left Behind.” In industrial Asian countries, extra instruction was provided to allow passing students to become better and to help better students become excellent. The strategy worked. The percentage of students reaching the highest level in proficiency in mathematics was twice as high in industrial Asian countries as in the United States.

**Opportunities and willingness to learn.** The intelligent person is, almost by definition, alert to opportunities to use ideas other than his or her own. The same is true of societies. The attitude “If it wasn’t invented here, it isn’t wanted” is a recipe for national ignorance. Historically, though, there have been examples of nations willfully refusing to listen to outside ideas. Modern Japan was cited as an example of the positive effects of developing national cognition. Four centuries ago, things were different. In the early 17th century, the ruling shogun decided that the nation should be isolated to avoid corrupting foreign influences. Japan sat out the industrial revolution, choking off access to a burst of cognitive artifacts developed in the West. The break was entirely due to national policy. The Japanese were aware of Western, innovative societies but chose not to listen to them. When Japan reopened itself in the 19th century, GCA bloomed.

In other nations, the development of GCA has been handicapped because the nations involved were simply “out of the loop” of cognitive innovation. For over 2,000 years, ideas have spread through the Eurasian continent along the Silk Road from China through the Mediterranean and into Europe via Venice and Genoa. Societies along the Silk Road and the trade routes of Europe were exposed to new ideas and artifacts; societies off the trade routes became cognitive backwaters. The people of the New Guinean highlands have a credible claim to being one of the world’s first agriculturists (Denham et al., 2003). The highlanders, isolated by mountains from the coast of an island that was itself isolated from the Asian mainland, stayed in the Stone Age until the 1930s (Diamond, 1997).

The history of the adoption of cognitive artifacts shows a similar sensitivity to geography. The rudiments of writing systems and mathematics spread throughout Eurasia. In the Americas, the jungles of northern South America and the isthmus of Panama cut off Incan and related civilizations in South America from the Mesoamerican civilizations to the north. Communication by sea never developed because neither South

---

**Fig. 3.** A: Students in industrialized Asian countries receive more extracurricular instruction in mathematics than do students in industrialized non-Asian countries. B: The Program for International Student Assessment scores of Asian students increase with increasing amounts of instruction, whereas the scores in non-Asian countries decrease as extracurricular instruction increases.

Effect could be because girls and women, as individuals, were more motivated to study in societies that offered them opportunities to use their education, or it could be that the societies, intending to use the cognitive resources women provided, were more interested in supporting women who wished to study. The two explanations are not mutually exclusive. From the national viewpoint, the results are what matters. Gender inequality results in an underdevelopment of STEM-relevant cognitive resources in half the population. That is not a good strategy for a country that wishes to occupy a leading position in the global economy of the 21st century.

A second contrast between Asian and European and North American countries also reflects differences in attitudes toward educational achievement. The upper panel of Figure 3
nor Central Americans had a seafaring culture. Even if they had, they would have had difficulty making contact by sea because the Mesoamericans were on the Caribbean side of the isthmus of Panama, while the Inca were on the Pacific side of South America.

The Inca had no writing system; the Maya and Aztec had writing systems that combined pictographic, logographic (symbol–word correspondence), and phonetic elements. The Inca had no formal mathematics as such, although they did have a system of accounting using knotted ropes (quipu). The Maya and Aztecs had base systems capable of representing any number and used these systems in astronomical calculations and possibly to determine land areas. And tragically for these civilizations, neither the Maya nor Aztecs were forced to develop very much in the way of military technology, especially tactics. The Maya and the Aztecs fell to Spaniards who had worse armor than they did, better swords, generally ineffective guns, and a tradition of fighting in tactically controlled formations, a cognitive artifact that stretched back to Greek times.  

Today communications are virtually worldwide, so every society has a chance to learn about the cognitive artifacts of others. They do not always take this chance. In 2003, there was an outbreak of polio in northern Nigeria. Modern medicine has an effective vaccine for this disease, but local leaders neither understood nor trusted Western medical practices. Rumors spread that the vaccine was loaded with the HIV/AIDS virus and that it was designed to sterilize young women, as a way of reducing the Muslim population in the country. Vaccination was initiated only after several years of negotiation and explanation between WHO, the central government of Nigeria, and local traditional and religious leaders. Meanwhile, the disease spread from Nigeria to several previously polio-free countries in Africa (Jegede, 2007).

The Nigerian case was particularly damaging, due to a number of special conditions. Less damaging, but still serious, denials of well-validated, modern concepts of medicine and public health occur in the I&PI world. In 2011, there were protest movements against the vaccination of schoolchildren in Canada, the United Kingdom, and the United States—if these movements became widespread they would have serious repercussions for the health, including the cognitive capacity, of individuals and the nation as a whole.

Modern refusals of obviously useful cognitive artifacts are not confined to the developing world, to health affairs, or to the actions of Luddite minorities in postindustrial nations. As of 2012, there is one contemporary nation that thinks of itself as a paragon of postindustrial economic and scientific innovation, yet refuses to adopt the metric system.

### Discussion

The analysis that has been presented here rests on three contentions: that intelligence and educational tests are legitimate measures of some of the cognitive skills required to succeed in the I&PI world, that national differences in such cognitive skills exist, and that the differences between national populations are largely due to environmental differences. There is a further subcontention that international differences in GCA are largely due to the same environmental variables that influence GCA within a nation. I have taken an intermediate position with regard to international differences in genetic potential. They might be important, or they might not. There is no compelling evidence either way.

There are two major objections to my argument: The first is that I have incorrectly evaluated the argument for genetic differences; the second is that the entire approach is based on a culturally biased view of cognition that is not applicable outside of the I&PI world. These are very different objections.

#### The genetic hypothesis (again)

Some psychologists have taken very strong views about the hypothesis that differences in GCA between racial–ethnic groups have a genetic origin. These views are directly relevant to discussions of international differences, as international differences in test scores are closely related to the racial and ethnic compositions of different countries. Two of the most vocal proponents of the genetic hypothesis have this to say: “Genetic and cultural factors carry the exact same weight in causing the mean Black-White difference in IQ as they do in causing individual differences in IQ, about 80% genetic–20% environmental by adulthood” (Rushton & Jensen, 2005, p. 279). An opponent of the genetic hypothesis has examined the evidence and concluded that “For the race differences in IQ, we can be confident that genes play no role at all” (Nisbett, 2009, p. 197).

Neither of these extreme statements can be justified. As discussed in detail above, there is no direct evidence for the Rushton and Jensen (2005) position. There are indirect arguments for their position, but the arguments are weak ones. Conversely, absent direct evidence from an analysis of gene frequencies, the Nisbett (2009) position is equally untenable. Strictly speaking, it amounts to an assertion of the null hypothesis. Psychologically, the Nisbett position is understandable and may be attractive to many. Showing that the evidence offered for a proposition is weak is often interpreted as evidence that the proposition, here the genetic hypothesis, is false. However that does not follow logically, so the Nisbett position is not a good example of rigorous scientific reasoning.

Virtually all the parties to this debate are professors, and by their nature, professors profess. Until direct tests of the genetic hypothesis become possible, the correct thing to do is to profess ignorance.

#### The cultural imperialism objection

The gist of this objection is that it is simply inappropriate to use cognitive tests developed for the I&PI world to predict the cognitive capacities of people or populations outside that
world. A related argument is that even within the I&PI world, these tests have primarily been developed to predict secularly defined material success, such as socioeconomic status or income. There are other aspects of life, such as commitment to family or living a life in harmony with some philosophic principle (including religions) that may be highly valued but are not predicted by cognitive test scores.

The first argument is directed at the tests themselves. The extent to which the cognitive tests developed within the framework of the I&PI world can predict success in endeavors outside that world is an open question. To the extent that the tests evaluate properties of reasoning that are required by all cultures, such as the ability to control attention, orient in space, and understand language, one would expect some prediction. To the extent that the tests evaluate aspects of reasoning specific to the I&PI culture, such as the use of syllogistic reasoning, they may not predict success in other cultures. There is also the issue that the testing paradigm itself, the “drop in from the sky” method of evaluation, may be foreign to many cultures. Any investigator operating outside the I&PI world has to be concerned about this, but how well the investigator has met the challenge can be answered only on a case-by-case basis.

The second argument is directed not at the tests but at the criterion measures used to validate tests. Within the I&PI world, the use of cognitive tests is validated by their ability to predict certain measures of secularly defined material success. These typically include income, measured socioeconomic status, and overt academic achievement, as evidenced by grades achieved and degrees earned. Other societies may not define success in this way. Therefore, it is arrogant to define the intelligence of a nation outside the industrial world by the extent to which its population possesses the skills needed to succeed in that world.

It is true that over the years various groups, ranging from Galton in the 19th century to the advocates of genetic hypotheses in the 21st, have used differences in test scores to infer differences in genetic potential. It is also true that in some sense even the use of terms like “developed” and “developing” or “traditional” compared with “modern” nations carries with it the implication that there is a natural course of development of human societies, and that individual nations can be placed at their position in the developmental path. Why should the skills and beliefs of a Masai goatherder in Kenya be regarded as more or less valuable than those of a computer programmer in Seattle?

This argument has much to recommend it as an abstract statement of the equality of human endeavors. (It also has weaknesses, but a discussion of these would take us too far afield.) For the limited purpose of this article, though, the objection misses the central point of the argument. There is no claim that the I&PI definition of intelligence is in some sense morally superior to the definitions that might be offered by other societies. The claim is a pragmatic one. I&PI societies dominate the world today. Barring a cataclysm such as thermonuclear war, they will continue to do so for the foreseeable future.

Virtually all the developing nations of the world are striving to obtain the benefits of the postindustrial world. The extent to which they will succeed will depend very much on their ability to develop and use those cognitive abilities that are essential in I&PI societies.

Conclusion

Nations do indeed differ in their cognitive resources. So do states within the United States (Pesta, McDaniel, & Bertsch, 2010), although the differences are not nearly as extreme as the international differences. These differences influence the economic, social, and physical well-being of the relevant society.

The thesis of this article is that national indicators of intelligence are indicators of national differences in the ability to use the cognitive artifacts that are required to participate in modern industrial and postindustrial societies. Differences in national capabilities to use cognitive artifacts are due to differences in the extent to which different nations provide techniques and institutions for the development of individual cognition. Many of the efforts to improve cognition interact, as in the case of nutrition and education. In other cases, efforts to improve cognition may conflict with short-term economic goals, such as the development of natural resources. I conjecture that, possibly outside of schooling, it may be harder to make social changes than to make changes in the physical environment.

Acknowledgments

This article is based on an invited address to the Association of Psychological Science (APS) in Washington, DC, in May 2001, on the occasion of the author’s receiving the James McKeen Cattell award for lifetime accomplishment. I thank APS for awarding me this unexpected honor. A lifetime award is almost always a team award, presented to the coach. I have been fortunate to work with a fine group of colleagues: undergraduates, graduates, postdoctoral and faculty associates, research technicians, and administrative associates. I am happy to accept this honor on behalf of the team.

Declaration of Conflicting Interests

The author declared that he had no conflicts of interest with respect to his authorship or the publication of this article.

Notes

1. The correlation is between the occupation’s GCA loading and the logarithm of median wages for an occupation, as listed by the Bureau of Labor Statistics, Income Census, 2011.
4. I wish I could claim credit for this phrase describing testing, but I cannot. I first heard it used by the psychometrician Robert Mislevy, then working at the Educational Testing Service.
5. The qualification “of comparable cost” is important. In academic settings, prior grades are often as good or slightly better indicators of success than are tests. In industrial settings, work samples are approximately as good indicators as test scores. Carefully structured interviews are also approximately as accurate predictors of future performance as are test scores. (Unstructured interviews are not even close to as accurate.) In academic settings, personnel screening occurs primarily at the college/university entry level. The use of grades presents a problem if applicants come from secondary schools that have varying standards. It is more expensive to obtain a work sample or conduct a structured interview than to give a test.

6. The predictive correlation is estimated by correcting the observed correlation between test scores and the performance of successful applicants in order to allow for restriction of range and unreliability of the test. A case can be made for correcting only for restriction of range. However, because conventional tests are highly reliable, the correction for unreliability is usually small.

7. Studies such as Terman and Oden’s (1959) and the SMPY have been criticized on the grounds that the “gifted” groups studied did not include any extremely distinguished contributors to society, such as a Nobel laureate. This criticism sets an impossibly high criterion for the success of the study. Consider the following rough calculation: The SMPY and Terman and Oden studies, combined, contained slightly less than 4,000 people. As of the end of 2011, 549 Nobel Prizes had been awarded (number downloaded from http://www.nobelprize.org/nobel_prizes/nobelprize_facts.html, January 2012). Suppose that an equal number of people made contributions as great as the laureates but, for some reason, did not receive the prize. Combining these numbers, we estimate that there were approximately 1,100 extremely distinguished contributors to society over the last 150 years. Now assume further that due to a variety of restrictions, such as cultural background, only the most favored 5% of the people alive during the period in which the studies were conducted came from a population that could include a distinguished contributor. (For instance, we disregard the rural populations of India and China, which alone constituted about 40% of the world’s population over this time period.) Call this 5% the eligible group. Assuming that roughly 10 billion people had adult lives during the past 150 years, the eligible group would contain approximately 500 million individuals, including 1,100 distinguished contributors. The probability of randomly choosing a distinguished contributor from the eligible group would be \(0.0000022 \times 4000\), just over two chances in a million. The probability of randomly choosing a distinguished contributor from the eligible group would be \(0.0000022\), just over two chances in a million. The eligibility of randomly chosen group of 4,000 members of the eligible population did not contain any distinguished contributors would be \((1 - 0.0000022)^{4000}\) — that is, .99.

8. Economic data for Japan and Nigeria obtained from the CIA World Factbook, January 2012.


10. The reference nations were Brazil, Ghana, India, Norway, Oman, and the United States.


12. Based on the CIA World Factbook country profiles as of 2011. The African nations of the Sudan, South Sudan, and Somalia have been omitted because at the time of writing, a substantial portion of their populations were refugees from civil wars.

13. Lynn and Vanhanen (2002, 2006) have made much lower estimates, but as Wicherts and his colleagues have pointed out (Wicherts, Dolan, Carlson, & Van Der Maas, 2010b; Wicherts, Dolan, & Van Der Maas, 2010), Lynn and Vanhanen used unsystematic selection methods that biased their estimates downward.

14. The data on GDP were provided by the CIA World Factbook (2011). The data for secondary education fees were provided by Keratu School Foundation, Keratu, Tanzania. This figure includes tuition (which is low) and numerous fees, which collectively add up to quite a bit.

15. The People’s Republic of China, currently the most populous country on earth, is a unique case. Regions of China, such as Shanghai, are industrialized and, as shown here, tend to have high test scores. China has not reported test scores for its large, less industrialized regions.

16. The Spanish conquistadores did not prevail solely because of their cognitive skills. The Amerindians were greatly weakened by the inadvertent introduction of smallpox, for which they had no immunity. However, the conquistadores won almost every battle that they fought, even though they were often heavily outnumbered. A major factor was their use of infantry tactics similar to modern close-order drill, which ensured that although they were outnumbered on the field, they outnumbered the enemy at the point of contact. Police today control large crowds in a similar way. The conquistadores also made use of superior writing and communication systems to report to and call for reinforcements from Spain. Our modern, secular society often underrates the influence of the Catholic Church as a motivating force that helped maintain morale during the arduous campaigns of colonial conquest. Tactics, writing, and religion are cognitive artifacts.

References


Lynn, R., & Harvey, J. (2008). The decline of the world’s IQ. *Intelligence, 38*, 112–120.


