American Psychological Association

Psychological Bulletin

0033-2909/19/$12.00

2019, Vol. 145, No. 3, 237–272
http://dx.doi.org/10.1037/bul0000184

Spearman's *g* Found in 31 Non-Western Nations: Strong Evidence That *g* Is a Universal Phenomenon

Russell T. Warne and Cassidy Burningham
Utah Valley University

Spearman's *g* is the name for the shared variance across a set of intercorrelating cognitive tasks. For some—but not all—theorists, *g* is defined as general intelligence. While *g* is robustly observed in Western populations, it is questionable whether *g* is manifested in cognitive data from other cultural groups. To test whether *g* is a cross-cultural phenomenon, we searched for correlation matrices or data files containing cognitive variables collected from individuals in non-Western, nonindustrialized nations. We subjected these data to exploratory factor analysis (EFA) using promax rotation and 2 modern methods of selecting the number of factors. Samples that produced more than 1 factor were then subjected to a second-order EFA using the same procedures and a Schmid-Leiman solution. Across 97 samples from 31 countries totaling 52,340 individuals, we found that a single factor emerged unambiguously from 71 samples (73.2%) and that 23 of the remaining 26 samples (88.5%) produced a single second-order factor. The first factor in the initial EFA explained an average of 45.9% of observed variable variance (SD = 12.9%), which is similar to what is seen in Western samples. One sample that produced multiple second-order factors only did so with 1 method of selecting the number of factors in the initial EFA; the alternate method of selecting the number of factors produced a single higher-order factor. Factor extraction in a higher-order EFA was not possible in 2 samples. These results show that *g* appears in many cultures and is likely a universal phenomenon in humans.

Public Significance Statement
This study shows that one conceptualization of intelligence—called Spearman's *g*—is present in over 90 samples from 31 non-Western, nonindustrialized nations. This means that intelligence is likely a universal trait in humans. Therefore, it is theoretically possible to conduct cross-cultural research on intelligence, though culturally appropriate tests are necessary for any such research.

Keywords: Spearman's *g*, cross-cultural psychology, general cognitive ability, human intelligence, factor analysis

For several decades the question of the nature of intelligence has been a matter of debate, with various scholars producing different definitions (see Sternberg & Detterman, 1986, for a compendium of definitions). One early prominent intelligence researcher stated that "An individual is intelligent in proportion as he is able to carry out abstract thinking" (Terman, 1921, p. 128, emphasis removed). The emphasis on abstract thought is apparent with other definitions throughout the 20th century (e.g., Freeman, 1923; Gottfredson, 1997a). However, other experts believe that intelligence consists of more than the capacity for abstract thought, such as Porteus (1965), who believed foresight and planning were important components of intelligence. Conversely, Cattell argued that intelligence consisted of fluid abilities—such as abstract reasoning—and crystalized abilities—such as learned knowledge of culturally relevant facts (Gustafsson & Wolff, 2015). Yet, others expand the term "intelligence" to include the entire breadth of cognitive skills,

This article was published Online First January 14, 2019.
Russell T. Warne and Cassidy Burningham, Department of Behavioral Science, Utah Valley University.
Cassidy Burningham is now at the College of Social Work, University of Utah.
We appreciate Salaheldin F. Bakhiet, Michael J. Boivin, Amelia Díaz, Juana Gómez-Benito, Tippawan Pongcharoen, Nicolas Ruffieux, Ola Sternäng, and Fons J. R. van de Vijver for their willingness to share their data with us in the spirit of open science and cooperation. We commend the foresight of the researchers who made their data or correlation matrices available publicly for future analyses. We acknowledge the help from two librarians at the Utah Valley University Fulton Library: Karen Sturtevant (for whom no interlibrary loan request is too difficult or obscure) and Emily Bullough (for helping us devise a systematic literature search procedure). We also greatly appreciate all the people who have collected the data we analyzed in this article, many of whom administered psychological tests under very difficult circumstances. A poster based on this study was presented on July 13, 2018, at the International Society for Intelligence Research in Edinburgh, Scotland. Study preregistration, materials, annotated data files, and factor analysis results are available at https://osf.io/tgq5y/.
Correspondence concerning this article should be addressed to Russell T. Warne, Department of Behavioral Science, Utah Valley University, 800 West University Parkway MC 115, Orem, UT 84058. E-mail: rwarne@uvu.edu

abilities, and acquired knowledge (e.g., Cleary, Humphreys, Kendrick, & Wesman, 1975; Conway & Kovacs, 2018; Halpern & Butler, 2018). Carroll (1993) proposed a compromise position among all of these viewpoints which postulated that intelligence was a general mental ability that was functionally related with other important narrower abilities.

These contending definitions lead some individuals to believe that intelligence is not definable and therefore the research on the topic is inherently flawed (e.g., Singham, 1995). Some have even used Boring's (1923, p. 35) circular definition that ". . . measurable intelligence is simply what the tests of intelligence test . . ." to claim that psychologists do not even know what intelligence is and that their efforts are hopelessly misguided (e.g., Richardson, 2002).

Cross-Cultural Definitions of Intelligence

One criticism that sometimes appears in the scholarly literature is that these definitions all originate from Western cultures, which may limit experts' perspectives on the nature of intelligence (Berry & Bennett, 1992). For example, Berry (1986), argued that ". . . I conceive of *intelligence*, as presently used in psychology to be [a] culture-bound, ethnocentric, and excessively narrow construct" (p. 35, emphasis in original). In accordance with this claim, it seems that many cultures have some definition of what intelligence is and what this means to people. As a result, some theorists contend that considering other cultural perspectives would broaden theoretical perspectives beyond the traditional definitions of intelligence (e.g., Sternberg & Grigorenko, 2001).

Non-Western cultural perspectives may not be precisely equal to one another, but there are some commonalities. Many non-Western groups value cognitive competence. For example, in India, people value intelligence and consider it an important aspect in their work and everyday life (Srivastava & Misra, 2001). But many non-Western cultures also believe that intelligence has a social and/or emotional component. Striving to accomplish a task is valued in India, East Africa, and elsewhere (Berry & Bennett, 1992; Jones, Rakes, & Landon, 2013; Srivastava & Misra, 2001), while in Zimbabwe, caring for and helping one's relatives before helping friends or strangers is an important social component of intelligence (Mpofu, 2004). Zimbabweans also value social competence, with a wider array of domains that high ability can show itself in, including lovemaking and witchcraft (Ngara & Porath, 2004). These "noncognitive" skills may be more important for success in a rural Zimbabwean context than the skills that Westerners see as important aspects of intelligence (Serpell & Jere-Folotiya, 2008). The native Cree peoples of Canada see traits like patience, deliberation, and persistence as part of cognitive competence (Berry & Bennett, 1992). Thus, these non-Western cultures see the cognitive components of intelligence as essential, but they do not limit their conceptualizations of intelligence to cognitive skills (Sternberg, 2003). Rather, these cultures also value emotional and social competence.

This mix of cognitive and social skills sometimes appears in the literature as "wisdom," which combines knowledge but also the ability to use that knowledge in prudent, moral, or socially useful ways. Such viewpoints have been found in Taiwan (Yang & Sternberg, 1997), Zimbabwe (Serpell & Jere-Folotiya, 2008), and First Nations individuals in Canada (Berry & Bennett, 1992), and elsewhere. This has led at least one prominent Western researcher to attempt to incorporate the concept of wisdom into theorizing about intelligence (Sternberg, 2003).

Adding to the complexity of defining intelligence across cultures is the fact that different cultures value different specific skills (Mpofu, Ntinda, & Oakland, 2012; Sternberg & Grigorenko, 2001; Vernon, 1969). For example, Berry (1966) found that Inuit examinees outperform rural Sierra Leonese individuals in spatial tasks, possibly because the Inuit environment and culture foster the development of spatial ability as a skill one needs to navigate the vast Arctic terrain. Likewise, Kathuria and Serpell (1998) found that Zambian children were largely unfamiliar with drawing with pencil and paper, thus making the draw-a-person test unsuitable for them. But Zambian children were much more experienced and adept at molding clay, an astute insight which resulted in the creation of the culturally appropriate Panga Munthu test of intelligence. This test requires a child to create a human figure using clay or plasticine. Because different cultures value different skills, some theorists have suggested that examinees from different cultures require different intelligence tests (e.g., Bernardoni, 1964; Sternberg & Grigorenko, 2001) or that intelligence tests are completely inappropriate for some cultures (e.g., Kwate, 2001; Ngara & Porath, 2004; Nsamenang, 2006) and that comparing scores across cultures is nonsensical (Vernon, 1965a).

Not Intelligence, but Spearman's *g*

This review of contending definitions of intelligence shows the futility of attempting to find agreement among experts—let alone among diverse cultures worldwide. Terms for "intelligence" or other cognitive abilities will often not translate directly across languages (Berry, 1966; Booth, 2002; Kathuria & Serpell, 1998). Even if a language's word can be translated as "intelligence," it may not be realistic to expect that both words will encompass the same skills and convey the exact same nuances.

Investigating cultural beliefs about intelligence may be mildly interesting from an anthropological perspective, but it sheds little light on the nature of intelligence. One undiscussed methodological problem in many studies of cultural perspectives on intelligence is the reliance on surveys of laymen to determine what people in a given culture believe about intelligence. This methodology says little about the actual nature of intelligence. The same logic that researchers use to argue that a folk belief regarding intelligence provides evidence of the nature of intelligence could also be used to argue that widespread cultural beliefs in elves, goblins, or angels provide evidence of the existence of supernatural beings. Survey methodology regarding layman beliefs cannot substitute for collecting data on individuals' performance on mental tasks.

Additionally, the emphasis that different cultures place on various cognitive skills shows that any definition of "intelligence" will be value-laden. Western and non-Western viewpoints of intelligence are intertwined with the skills and knowledge that individuals in a culture value as being important for survival, social harmony, and/or individual success. Differing values will inevitably lead to contending definitions of intelligence, and there is no objective way to determine which values should have priority.

Because "intelligence" can never be defined in a manner that satisfies experts and laymen from many cultures, we believe that individuals who wish to understand the cross-cultural nature of mental abilities should focus their attention instead on Spearman's *g*. For over a century (since Spearman, 1904) some psychologists have postulated that a general mental ability factor can account for approximately half of the variance shared among intercorrelating cognitive test scores (Canivez & Watkins, 2010). This construct, rather than being verbally defined, is a statistical observation, which frees scholars from the need to find an exact definition of "intelligence." Indeed, Meehl (2006, p. 435) recognized the value of *g* over "intelligence," when he stated "Verbal definitions of the intelligence concept have never been adequate or commanded consensus," and that works focusing on *g* instead ". . . essentially solve the problem" of an inadequate definition of intelligence.

From early in its history, *g* has had the potential for cross-cultural research. Spearman (1927, pp. 164–167) saw *g* as an ability that had three characteristics: (a) the "apprehension of one's own experience," (b) the "eduction of relations," and (c) the "eduction of correlates." In modern terms, these characteristics would be the ability to (a) engage in metacognition and self-reflection, (b) understand how items and concepts are related to one another, and (c) generate new ideas or concepts on the basis of existing an idea and a knowledge of a relationship. Spearman believed that individual differences would exist in such an ability and in self-reflection and the ability to comprehend relationships or use relationships to generate new ideas. Moreover, he saw these characteristics as being universal because all humans self-reflect (e.g., via memory recall of their actions) and understanding relationships—like cause and effect, or before and after. However, Spearman based his ideas entirely on theory, and did not gather any cross-cultural information to test his ideas. Later generations of researchers, however, have found that *g* does appear in different countries (e.g., Carroll, 1993) or used *g* in their cross-cultural investigations of individual differences in cognitive ability (e.g., Jensen, 1998).

Another benefit with focusing on the *g* is that it is easy to measure. As long as a task is at least somewhat cognitive in nature, it will be at least a partial measure of individuals' general cognitive ability. Spearman (1927, pp. 197–198) called this idea the "indifference of the indicator." According to this theory, many different tasks can measure *g*, and these tasks do not need to resemble traditional intelligence test items at all (Cucina & Howardson, 2017; Hunt, 2011). Indeed, some people have theorized that many everyday tasks in individuals' lives require at least some cognitive effort; these daily tasks may function as measures of *g* (Gottfredson, 1997b; Lubinski, 2004), which would explain why so many life outcomes correlate with individuals' level of general cognitive ability (see Herrnstein & Murray, 1996; Hunt, 2011; Jensen, 1998; or Warne, 2016, for examples). Because of the indifference of the indicator, it is reasonable to expect that *g* can emerge from test data regardless of an individual's cultural background—so long as the examinee understands the task and stimuli and that the task requires cognitive work or effort.

Finally, *g* has the benefit of being much more difficult to train than narrower cognitive abilities or some of the traits are sometimes considered part of "intelligence" throughout the world. While training for specific mental abilities (e.g., working memory capacity, executive functioning) is feasible, *g* seems to be much more resistant to training, at least in industrialized societies (Melby-Lervåg, Redick, & Hulme, 2016; Protzko, 2017; Schwaighofer, Fischer, & Bühner, 2015).[1] Spearman's *g* is also highly heritable in Western nations (Plomin, 2001; Plomin, DeFries, Knopik, & Neiderhiser, 2016). These two characteristics make *g* likely a robust phenomenon that is less susceptible to cultural differences than many other psychological traits—including traits that some may consider to be part of "intelligence."

Possible Cross-Cultural Nature of *g*

Spearman's *g* originated from a British researcher working with British subjects, which justifiably could lead some individuals to be skeptical about whether *g* would exist in other cultures. However, there are strong theoretical reasons to believe that a general cognitive ability would be part of the cognitive architecture of all human populations. First, *g* has been found to correlate with a wide variety of biological variables, including brain size (Pietschnig, Penke, Wicherts, Zeiler, & Voracek, 2015; Rushton & Ankney, 2009), white matter tract integrity (Penke et al., 2012), allele frequency (Hill et al., 2016), pupil size (Tsukahara, Harrison, & Engle, 2016), and myopia (Lubinski & Humphreys, 1992). Although most of this research has been conducted on Western populations, it would be extremely surprising for *g* scores to correlate with these fundamentally biological variables only within a genetically European population because the basic nervous system architecture of all humans is the same (Spitz, 1978).

A second reason why *g* would be expected in data from non-Western populations comes from ethology. Researchers studying the cognitive abilities of animals have identified a general factor in cognitive data from many mammal species, including nonhuman primates (Fernandes, Woodley, & te Nijenhuis, 2014; Herndon, Moss, Rosene, & Killiany, 1997; Hopkins, Russell, & Schaeffer, 2014; Matzel & Sauce, 2017), dogs (Arden & Adams, 2016), and rats and mice (Anderson, 1993; Galsworthy, Paya-Cano, Monleón, & Plomin, 2002; Matzel & Sauce, 2017). The most parsimonious explanation is that general cognitive ability evolved early in the evolutionary history of mammals, long before primates diverged from other mammalian species, and that these living mammalian species retained a general cognitive ability to the present day. Thus, a general cognitive ability would likely be a ubiquitous trait

[1] It is valuable to recognize the difference between *g* and IQ in regards to training. The two are not synonymous; IQ is a measure of the construct of *g* (Haier, 2017). Because measures of *g* (e.g., intelligence tests, cognitive tasks) will inevitably include tasks and abilities that are not *g*—such as language comprehension or explicitly taught skills, like knowledge of geometric shapes—IQ variance is made up of *g* and non-*g* components. Undoubtedly, IQ scores can be increased through training and educational programs. The Flynn Effect—where IQ scores increased worldwide throughout the 20th century—is an example of IQ score increases due to environmental changes (Flynn, 1987). Research into the Flynn Effect indicates that IQ score increases are due to changes into the non-*g* sources of IQ and not *g* itself (Woodley, te Nijenhuis, Must, & Must, 2014; Woodley of Menie, Piffer, Peñaherrera, & Rindermann, 2016). This also explains why Raven's progressive matrices, a heavily *g*-loaded test, can have a strong Flynn Effect: the non-*g* component is highly susceptible to training, resulting in higher scores for more recently born examinees (Armstrong et al., 2016; Armstrong & Woodley, 2014). Likewise, compulsory education can increase a population's average IQ score (Ritchie & Tucker-Drob, 2018), but there is no evidence that this increase in IQ is equivalent to an increase in *g*. Indeed, given the breadth of schooling and the specificity of many curricula, it is likely that schooling's impact on narrow, non-*g* abilities outstrips its impact on *g*.

in mammals—including humans from all cultures. Given the phylogenetic relationships that develop from all mammals having a common evolutionary ancestor (Hodos & Campbell, 1969), this explanation of why *g* appears in many mammalian species is plausible.

Alternatively, one could postulate that a general cognitive ability is a Western trait but not a universal trait among humans, but this would require an evolutionary model where this general ability evolved several times independently throughout the mammalian clade, including separately in the ancestors of Europeans *after* they migrated out of Africa and separated from other human groups. Such a model requires (a) a great deal of convergent evolution to occur across species occupying widely divergent environmental niches and (b) an incredibly rapid development of a general cognitive ability while the ancestors of Europeans were under extremely strong selection pressures that other humans did not experience (but other mammal species or their ancestors would have experienced at other times). We find the more parsimonious model of an evolutionary origin of the general cognitive ability in the early stages of mammalian development to be the more plausible one, and thus we believe that it is reasonable to expect a general cognitive ability to be a universal human trait.

Purpose of Study

Although we can generate plausible reasons why *g* would be a universal phenomenon in humans, empirical data are necessary to determine whether these theories are supported by evidence. The purpose of this study is to investigate the presence of Spearman's *g* in data from human populations that would be least likely to demonstrate it if *g* were a cultural artifact: non-Western cultures in nonindustrialized nations. As Booth (2002, p. 378) stated "If there are universals, they must be discovered," and for the reasons mentioned above, we think it is plausible that *g* may be a universal property of cognitive data in humans—whether or not a cultural group describes "intelligence" in a way similar to Western definitions. While there have been scattered reports of *g* in nations throughout the world (e.g., Church & Katigbak, 1987; Claeys, 1972; Guthrie, 1963; Miron, 1975; Proctor, Kranzler, Rosenbloom, Martinez, & Guevara-Aguire, 2000; Rothhammer & Llop, 1976; Vernon, 1961, 1965b), in this study we aim to systematically compile as many data sets as possible from non-Western populations in developing nations and determine whether each one produces a general cognitive ability factor using a uniform factor analysis methodology. In this way we hope to put *g* theory to a strongly falsifiable empirical test and to test whether *g* theory has limits in its applicability across human cultures.

In addition to being a robust test of theory, we also view our work as a replication in three forms. The first is that we generally followed Carroll's (1993) procedures in his landmark report of his factor analysis procedures of 477 data sets from 19 countries, though with some updates to modernize the methodology (see below). The second form that our work is a replication is that we reanalyzed previously published data, which is ". . . a form of *replication*—a procedure recognized as critical in all scientific investigation" (Carroll, 1993, p. 80). This is because subjecting the same data to a different, improved methodological procedure is a test of whether the original results are robust. Finally, our study may also be considered a replication because it applied the same methodology to many independent data sets in order to determine whether the results are similar. Given the crisis of confidence in the replicability of many published psychological studies (Camerer et al., 2018; Open Science Collaboration, 2015), our three-pronged replication of prior research on the structure of cognitive abilities is a valuable endeavor.

Method

Search Procedure

Before beginning our search for data we defined a priori the populations that would qualify for the study. We defined "non-Western" nations as countries with fewer than 50% of their population classified as White or European, based on statistics from the World Factbook (Central Intelligence Agency, 2017). We defined "nonindustrialized" nations as countries that the United Nations (2018) had ever labeled as a "least developed country" or which were currently labeled "low middle income countries." The final list of 81 nations that were both non-Western and nonindustrialized are listed in Table 1 and shown on the map in Figure 1.

After we identified the qualifying countries, we preregistered our search and analysis procedures through aspredicted.org (see https://osf.io/tgq5y/ for preregistration information). Following our preregistration protocol, we searched for the terms (cognitive OR intelligence) AND *name of country* in PsycINFO, PsycARTICLES, ScienceDirect, JSTOR, AnthroSource, Web of Science, and a multidatabase search of our institution's library. We also searched for samples by investigating the reference lists of Carroll's (1993) book and the samples listed in David Becker's update of Richard Lynn's database of international IQ averages (then-current version available at https://drive.google.com/file/d/0B3c4TxciNeJZNUVaZ0VhZ3dUMkk/view).

For inclusion in our study, data sets must contain scores from at least four cognitive tasks or tests. Sample members were also required to be predominantly or entirely nonclinical individuals, and samples that consisted of individuals who were selected on the basis of a disability, physical disease, or psychological diagnosis were not included. The raw data or a correlation matrix and sample size were also required for inclusion in the study.

Analysis

After our search, we performed an exploratory factor analysis (EFA) on each correlation matrix or dataset found using principal axis extraction and promax rotation. We chose this nonorthogonal rotation method because an orthogonal rotation method would require cognitive abilities to be perfectly uncorrelated, which we thought was unrealistic because scores on mental tasks are frequently positively correlated—a finding which has been apparent in psychological research for many years (e.g., Carroll, 1993; Jensen, 1998; Spearman, 1904; Yerkes, 1921).

One of the most important—and yet most subjective—decisions in an EFA is the method for retaining the number of factors. We used two methods for making this decision: (a) Larsen and Warne's (2010) modification of the Guttman rule of factor selection, and (b) Velicer's (1976) minimum average partial (MAP) method. The former rule retains a factor in a dataset if the entire 95% confidence interval around the factor's eigenvalue is greater

Table 1
Non-Western, Nonindustrialized Countries

Afghanistan	Guyana	Philippines
Angola	Haiti	Rwanda
Bangladesh	Honduras	Saint Lucia
Benin	**India**	Saint Vincent and the Grenadines
Bhutan	**Indonesia**	Samoa
Bolivia	**Kenya**	Sao Tome and Principe
Botswana	Kiribati	Senegal
Burkina Faso	Laos	**Sierra Leone**
Burundi	Lesotho	Solomon Islands
Cabo Verde	Liberia	Somalia
Cambodia	**Madagascar**	South Sudan
Cameroon	**Malawi**	Sri Lanka
Central African Republic	Maldives	**Sudan**
Congo, Democratic Republic of the	Mali	Swaziland
Congo, Republic of the	Marshall Islands	Syria
Côte d'Ivoire	Mauritania	**Tanzania**
Djibouti	Mongolia	**Thailand**
Egypt	**Morocco**	Timor-Leste
El Salvador	Mozambique	Togo
Equatorial Guinea	Myanmar	Tonga
Eritrea	**Nepal**	Tuvalu
Ethiopia	Nicaragua	**Uganda**
Gambia	Niger	Vanuatu
Ghana	Nigeria	Vietnam
Guatemala	**Pakistan**	Yemen
Guinea	**Papua New Guinea**	**Zambia**
Guinea-Bissau	**Paraguay**	**Zimbabwe**

Note. Bold type indicates that the country was the origin of a dataset in the study. Eight countries did not have enough information to be classified: Antigua and Barbuda, Dominica, Grenada, Micronesia, Nauru, North Korea, Palau, and Saint Kitts and Nevis. These nations were not investigated in this study. Gauvain and Munroe (2009) reported an international dataset that also included individuals from Belize and American Samoa.

than 1.0—an attempt to compensate for the tendency for the Guttman rule to frequently identify many more factors than a dataset really contains. The MAP method identifies the number of factors to retain in the dataset by finding the factor at which the average of the squared partial correlations is at its smallest value after first partialing out the extracted factor(s) (Velicer, 1976). We used SPSS syntax from O'Connor (2000) to execute the MAP method. Prior research on methods for selecting the number of factors in an EFA indicates that the MAP method is generally the most accurate and that all other methods—including the modified Guttman rule—select too many factors (Fabrigar, Wegener, MacCallum, & Strahan, 1999; Hakstian, Rogers, & Cattell, 1982; van der Eijk & Rose, 2015; Velicer, Eaton, & Fava, 2000; Warne & Larsen, 2014; Zwick & Velicer, 1982, 1986). Still, no selection method is always accurate, so the best practice is to use multiple methods of selecting the number of factors in a dataset (Thompson, 2004).

When more than one factor was obtained in either method, the factor scores were subjected to a second-order EFA to determine whether the factor scores produced a general factor in order to follow the procedure established by Carroll (1993) in his similar reanalysis of cognitive data sets. This second-order EFA also used principal axis extraction and promax rotation and the same two methods for selecting the number of factors, along with a Schmid-Leiman solution to interpret the results (Schmid & Leiman, 1957; Wolff & Preising, 2005). We performed this second-order factor analysis because identifying multiple factors in a dataset is not sufficient evidence against *g* because factors may themselves be correlated strongly enough that their relationships may be explained by a higher-order *g* factor. Thus, in cases where multiple factors were present, a second-order EFA was necessary to test whether the multiple factors still were components of *g* or whether they were sufficiently separate for us to reach the conclusion that *g* was absent from the dataset. For all factor analyses, the percentage of variance explained by the first unrotated factor (another definition of *g*) was also recorded. We decided that a sample would be labeled as producing *g* if either (a) the initial EFA produced just one factor, or (b) the second-order EFA produced a single factor.

In our preregistration protocol, we specified that we would attempt an exploratory investigation of data sets that did not produce *g* in order to ascertain why *g* was not present. We believe that perhaps there could be a common characteristic of these data sets (e.g., from the same part of the world, same age group, a broader definition of cognitive variables). As will be apparent in the Results section, this exploratory investigation was not needed.

Deviations From Preregistration Protocols and *Ad Hoc* Decisions

For the sake of transparency, we find it important to explicitly state deviations from our preregistration protocol. First, in our preregistration, we stated that we would search for (cognitive OR intelligence) AND the name of a continent or population. However, searching for a continent was not feasible in finding data sets.

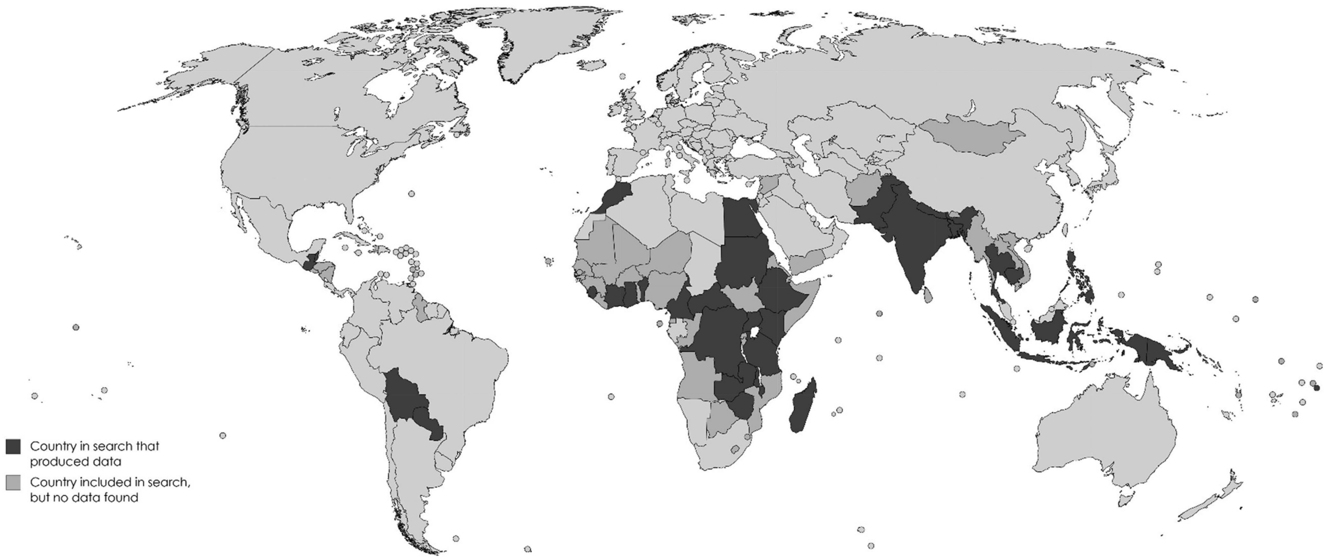

Figure 1. Map of countries included in the search. Countries in gray were included in the data search and are listed in Table 1. The darker gray countries are represented in the study, while the lighter gray nations were included in the search but did not produce any data.

We also had difficulty generating a list of population groups (e.g., ethnic groups, tribal groups) that would be useful for our search procedures.

Although not part of our preregistration, we expanded our search procedures by examining the reference lists of every article that contained usable data for our study and of five articles about intelligence levels in developing nations (Lynn, 2010; Rindermann, 2013; Wicherts, Dolan, Carlson, & van der Maas, 2010a, 2010b; Wicherts, Dolan, & van der Maas, 2010) and Lynn's (2015) compilation of worldwide IQ test results. We also found that searching for specific intelligence test names, coupled with the terms "cognitive" or "intelligence" and a country name (e.g., *Wechsler intelligence Ghana*) in Google Scholar was productive. We also examined the entire archives of four journals that specialize in differential psychology and intelligence research (*Intelligence*, *Personality and Individual Differences*, the *Journal of Intelligence*, and *Mankind Quarterly*) to search for samples from the countries in Table 1. Finally, we contacted via e-mail two groups of researchers whom we believed would have access to relevant data. The first was leading researchers in the field of cross-cultural cognitive testing or testing in developing nations. The second group consisted of corresponding authors of articles in which it was clear that the data we desired had been collected but not reported as a correlation matrix. The list of authors who responded by sharing their raw data or an appropriate matrix is in our acknowledgments.

Another deviation from the preregistration protocol was in the decision method for the number of factors in a dataset. This is one of the most important decisions in an EFA (Thompson, 2004; Warne & Larsen, 2014), and these decisions can make *g* artificially easier or harder to identify. In addition to the two methods we used, we also originally intended to use parallel analysis (Horn, 1965), which simulation studies have shown is one of the more accurate methods of identifying the number of factors in a dataset (e.g., Warne & Larsen, 2014; Zwick & Velicer, 1986), which was one of the factor selection procedures Carroll (1993) used. We intended to use the computer syntax from O'Connor (2000) to conduct parallel analysis. However, we discovered that when applied to correlation matrices (instead of raw data), parallel analysis greatly overfactored and produced results that were not realistic. (The results indicated that every dataset contained more factors than would be suggested even by the original Guttman rule, which is known to regularly overfactor.) This overfactoring was because parallel analysis randomly shuffles variable scores within a dataset to produce a new dataset of random data with the same mean and standard deviation for all variables. Eigenvalues for this new dataset are compared with the original dataset's eigenvalues, and all factors with eigenvalues in the original dataset that exceed the random data's corresponding factors' eigenvalues are retained (Thompson, 2004). However, in a correlation matrix, the raw scores are unavailable to perform an accurate parallel analysis. Therefore, we eliminated the use of parallel analysis in determining the number of factors in the data sets.

As we found data sets, some unforeseen situations developed that required *ad hoc* decisions. For several data sets, authors used pairwise deletion to handle missing data, which resulted in an inconsistent sample size within a correlation matrix. Most authors (e.g., Berry et al., 1986; Humble & Dixon, 2017) reported the maximum sample size, and we used the sample size for our analyses. Another frequent situation arose when deciding which variables in a correlation matrix were "cognitive" variables. We decided that demographic variables—including educational attainment level—were not cognitive, nor were personality variables or measures of opinions or attitudes. These variables were excluded from analyses. In ambiguous situations we contacted the corresponding author to ask whether they considered a variable to be "cognitive" in nature. When this was not possible or feasible, we erred on the side of caution and included a variable because we wanted to maximize the chances of rejecting *g* and finding more than one general cognitive factor. This was our strategy because we wanted to subject *g* theory to the strongest falsification test possible.

We also decided that for inclusion in the study, a variable must be a measure of sample members' mental abilities, and not a rating from an informant, such as a peer, parent, or teacher (e.g., Grigorenko et al., 2001)—unless the rating was the result of a direct observation of the individual. Additionally, we occasionally found a dataset where examinees were given cognitive measures at multiple times, such as at various points in a longitudinal study or before and after an intervention (e.g., Bangirana, Sikorskii, Giordani, Nakasujja, & Boivin, 2015; Sternberg et al., 2002). In these situations, we chose to use data from the earliest measurement occasion possible. Finally, some samples consisted of data from multiple countries; usually all of the countries were on the list in Table 1, but one dataset (Gauvain & Munroe, 2009) included data from Belize and American Samoa. We made an *ad hoc* decision that this sample qualified for our study because the Belizean individuals were from an Afro-Caribbean community that spoke a native language and that Belize was impoverished enough when the data were collected (the late 1970s) to have been included on our list of nations in Table 1. The American Samoan sample would have been too economically prosperous for inclusion in our study had its data been reported alone, but these individuals were only 25% of sample members.

Most data that we found in our search was in the form of correlation matrices. However, raw data were available occasionally—a situation we had not anticipated. When possible we used multiple imputation with 20 imputations to handle missing data. For two data sets (Condon & Revelle, 2016), this produced unrealistic answers (four nearly uncorrelated factors produced from four observed variables after imputation, but the observed variables were moderately correlated using listwise deletion), probably because the variables were average scores of items that examinees had received in an item sampling procedure. For these two data sets we used listwise deletion to handle missing data. The same data source also produced data sets from other countries listed in Table 1. We chose to include data from Condon and Revelle (2016) only if there were 150 individuals from the same nation after listwise deletion procedures. This sample size was chosen *a priori* based on a recommendation from Thompson (2004, p. 24) regarding minimum sample sizes for EFA.

Finally, in seven studies (Berry et al., 1986; Bhatia, 1955; Gichuhi, 1999; Lean & Clements, 1981; Mahmood, 2013; Sternberg et al., 2001; Veena, 2014), there were two correlation matrices presented, and the "correct" matrix to analyze was not clear. For example, this could happen if one matrix had a variable partialed out (e.g., age) or if two matrices included a different set of variables. For these studies, we chose to analyze both matrices and present both sets of results. Six pairs of matrices produced the same number of factors and very similar percentages of variance in the first identified factor, while the seventh (Gichuhi, 1999) produced a slightly different number of factors, though the strength of the first factor was almost exactly the same.

Results

Our search produced 99 samples from 31 countries with a total of 52,839 individuals. Two data sets were found to report data from the same examinees (Grigorenko et al., 2001; Sternberg et al., 2001). This information was not explicitly stated in either article, but the sample and community description makes it clear that the participants of these studies are the same people, though the sample sizes differ slightly (*n*s = 85 and 86). However, this redundancy did not produce any analysis problems because the correlation matrix in the Grigorenko et al. (2001) article was not positive definite. Because this matrix was not analyzable we eliminated it from further consideration. Another correlation matrix (Brimble, 1963) was not positive definite, and so we also eliminated it from further analysis. Elimination of these two data sets resulted in an aggregate total of 97 samples from 31 countries totaling 52,340 individuals. The countries that the samples originated from are marked in bold on Table 1 and shown in dark gray in Figure 1.

The mean sample size of the remaining data sets was 539.6 (*SD* = 1,574.5). The large standard deviation in relationship to the mean is indicative of the noticeably positively skewed distribution of sample sizes, a finding supported by the much smaller median of 170 and skewness value of 6.297. There were 16,559 females (33.1%), 25,431 males (48.6%), and 10,350 individuals whose gender was unreported (19.8%). The majority of samples—62 of 97 samples (63.9%)—consisted entirely or predominantly of individuals below 18. Most of the remaining samples contained entirely or predominantly adults (32 data sets, 33.0%), and the remaining 3 datasets (3.1%) had an unknown age range or an unknown mix of adults and children). The samples span nearly the entire range of life span development, from age 2 to elderly individuals.

Table 2 contains a description of the data sets we found in our literature search. Thirty countries from Table 1 are represented in the data sets, plus Belize and the territory of American Samoa. The data sets come from every inhabited continent and region of the world, except Europe (eliminated in advance because of its Western cultures) and Australia (eliminated because of its economic prosperity and Western population and culture).

Table 3 displays the number of factors that were identified in each dataset when using the modified Guttman procedure and MAP. There were 75 (77.3%) data sets that were found to have one factor (presumably *g*) when using the modified Guttman procedure to determine the number of factors, and 82 (84.5%) data sets had one factor when using MAP. A total of 71 data sets (73.2%) had only one factor identified via both methods, meaning that 26 data sets had multiple factors when using either MAP or the modified Guttman procedure (26.8%).[2]

Among the 26 data sets that produced multiple factors with either the modified Guttman rule or MAP, almost all of them produced one second-order factor, which would presumably represent *g*. Only one dataset produced more than one second-order factor (Grigorenko, Ngorosho, Jukes, & Bundy, 2006). This dataset produced two first-order factors when using MAP and three

[2] The website https://osf.io/tgq5y/ includes of repository of our data analysis syntax files so that readers can verify our results. These files are annotated for the reader's convenience.

Table 2
Dataset Information and Factor Analysis Results

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
1	Abdelhamid, Gómez-Benito, Abdeltawwab, Bakr, and Kazem (2017)	Egypt	204	Range: 18–24 (M = 20.65, SD = 1.71)	128 (62.7%) females and 76 (37.3%) males	WAIS-IV adapted into Arabic	Correlation matrix obtained via personal communication (J. Gómez-Benito, January 17, 2018).
2	Agrawal and Kumar (1993, p. 288)	India	40	Range: 20–65	40 females (100.0%)	SPM, WAIS forward digit span subtest, WAIS backward digit span subtest, six reaction time tasks	
3	Agrawal and Kumar (1993, p. 288)	India	50	Range: 20–80	50 males (100.0%)	SPM, digit span, backward digit span, six reaction time tasks	
4	Alaraudanjoki (2003, p. 39, Table 9)	Nepal	150	Range: 10–14	71 (47.3%) females and 79 (52.7%) males	Bender visual-motor Gestalt test; WISC-R arithmetic, digit span, backward digit span subtest; and word fluency test	Subjects were all child laborers.
5	Alaraudanjoki (2003, p. 39, Table 9)	Nepal	104	Range: 10–14	47 (45.2%) females and 57 (54.8%) males	Bender visual-motor Gestalt test; WISC-R arithmetic, digit span, backward digit span subtest; and word fluency test	Subjects were all school children.
6	Attallah, Ahmed, and Meisenberg (2014, p. 158)	Sudan	262	Range: 6–16 (M = 8.9, SD = .94)	134 (51.1%) females, 128 males (48.9%)	WISC-III translated into Arabic and adapted for a Sudanese context	Subjects were all gifted children.
7	Bakhiet et al. (2017)	Sudan	1,214	M = 8.5, SD = .5	658 (54.2%) females, 556 males (45.8%)	WISC-III translated into Arabic and adapted for a Sudanese context	Correlation matrix obtained via personal communication (S. Bakhiet, January 31, 2018). Sample was applicants from Khartoum applying for gifted schools.

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
8	Bangirana, Sikorskii, Giordani, Nakasujja, and Boivin (2015)	Uganda	90	Range = 1.5–10.8 (M = 4.26, SD = 1.90)	47 (52.2%) females and 43 (47.8%) males	CogState detection task speed, identification task speed, One Card Learning task accuracy, One Back Memory task accuracy, Groton Maze Chase Test moves per second, and Groton Maze Learning Test errors subtest and Test of Variables of Attention Visual response time omission errors (total %), Visual commission errors (total %), Visual response time total, Visual response time variability total, visual D prime total subtests.	Data obtained via personal communication (M. Boivin, February 20, 2018). Sample size differs from published article because this analysis includes only the 90 healthy community control individuals who had not experienced malarial infections. KABC-II subtests and TOVA auditory subtests in dataset, but these were not included because there were too much missing data (>75%) for multiple imputation.
9	Bau and Dyck (1992, p. 33)	Papua New Guinea	195	Not reported	195 (100.0%) males	Pacific Reading Comprehension Test, Pacific Word Knowledge Test, Numerical Ability Test (two subtests), Speed and Accuracy Test (two subtests), Pacific Reasoning Series Test, math achievement, and science achievement	Subjects were military officers.
10	Belay (2006, p. 68)	Ethiopia	202	M = 30.8, SD = 7.8	202 (100.0%) females	KABC-II sequential processing subtest, KABC-II simultaneous processing subtest, KABC-II planning subtest, CPM	Women lived in three rural communities in southern Ethiopia.
11	Berry et al. (1986)[a]	Central African Republic	391	Unknown	Unknown	African Embedded Figures Test (two tasks), verbal test, and visual discrimination test speed and score.	Sample members were members of the Biaka, Bangandu, and Gbanu peoples.
12	Bhatia (1955, p. 48)	India	100	Unknown	100 (100.0%) males	1937 Stanford-Binet, Kohs block design test, Passalong, Patterns, Memory, and Pictures	Subjects were English-speaking boys at a private school in Allahabad where instruction was in English. An unknown proportion were multiracial (English and Indian).

(*table continues*)

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
13	Bogale et al. (2013, p. 278)	Ethiopia	100	$M = 28.7$, $SD = 5.6$	100 (100.0%) females	CPM, KABC-II short-term memory, visual processing, and pattern recognition indices.	Researchers seemed to have used pairwise deletion; *n* is maximum *n*. Women lived in rural communities in southern Ethiopia.
14	Brimble (1963, p. 33)	Zambia	499	Range = 12.5–14	499 (100.0%) males	Always has, picture classification, absurdities, analogies, sequence, and reverse similarities	Data collected in urban schools in Lusaka and Kabwe (then called Broken Hill). Matrix is not positive definite.
15	Bulatao and Reyes-Juan (1968, p. 44)	Philippines	234	Range = 13.4–18.3 ($M = 15.2$, $SD = .7$)	234 (100.0%) males	21 variables of perceptual ability, cognitive ability educational attainment, and creativity (the age, educational attainment, and "seatmate" rating variables were not included)	Students were enrolled in high school students in Quezon City.
16	Church, Katigbak, and Almario-Velazco (1985, p. 326)	Philippines	79	Unknown (most between ages 5–7)	37 (46.8%) females, 37 (46.8%) males, and 5 (6.3%) unknown	Picture identification, erroneous pictures, concept formation, picture construction, vocabulary, barrio information and comprehension, and WISC block design	Children were enrolled in a rural school in Bulacnin. Test is administered in Tagalog and is an adaptation of Western intelligence test item formats.
17	Conant et al. (1999, p. 380)	Democratic Republic of the Congo	84	Range: 6.1–8.5 ($M = 7.6$, $SD = .6$)	42 (50.0%) females and 42 (50.0%) males	K-ABC subtests	Subjects were children at a rural private Christian school.
18	Conant et al. (1999, p. 380)	Democratic Republic of the Congo	55	Range: 8.6–12.8 ($M = 9.7$, $SD = 1.1$)	24 (43.6%) females and 31 (56.4%) males	K-ABC subtests	Subjects were children at a rural private Christian school.
19	Condon and Revelle (2016)	India	589	Range = 14–67 ($M = 26.84$, $SD = 7.87$)	227 (38.5%) females and 362 (61.5%) males	3D rotation, letter and number series, matrix reasoning, and verbal reasoning subtests	Data collected as part of the International Cognitive Ability Resource. Sample members who did not answer at least one item on each subtest were removed from the dataset.
20	Condon and Revelle (2016)	Philippines	601	Range = 14–60 ($M = 22.33$, $SD = 7.34$)	441 (73.4%) females and 160 (26.6%) males	3D rotation, letter and number series, matrix reasoning, and verbal reasoning subtests	Data collected as part of the International Cognitive Ability Resource. Sample members who did not answer at least one item on each subtest were removed from the dataset.
21	Dasen (1984, p. 420)	Côte d'Ivoire	47	Range = 7.92–9.0	24 (51.1%) females and 23 (48.9%) males	Five Piagetian conservation tasks, six basic logic tasks, and spatial ability tasks	

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
22	Dawson (1967a, p. 120)	Sierra Leone	99	Range: 20–24	99 (100.0%) males	SL2 group intelligence test, Kohs block test, embedded figures test, and 3D image perception	Examinees were skilled workers.
23	Dawson (1967b, p. 179)	Sierra Leone	51	Range: 20–24	51 (100.0%) males	SL2 group intelligence test, Kohs block test, embedded figures test, and 3D image perception	Subjects were members of the Temne tribe.
24	Díaz, Sellami, Infanzón, Lanzón, and Lynn (2012)	Morocco	202	Range: 18–65 (M = 26.77)	110 (54.5%) females and 92 (45.5%) males	SPM, Gestalt continuation, verbal-analytic reasoning, and visuospatial ability	Correlation matrix obtained via personal communication (A. Díaz, January 18, 2018). Subjects were sampled from university residence halls, both residents and employees. 57% were students.
25	Elwan (1996, p. 104)	Egypt	172	Range: 6.3–12.2 (M = 9.3, SD = 1.7)	90 (52.3%) females and 82 (47.7%)	WISC-R similarities, WISC-R vocabulary, WISC-R block design, WISC-R picture completion, K-ABC Gestalt closure, K-ABC spatial memory, number recall, and hand movements subtests.	Subjects attended two public schools in Cairo.
26	Engle, Klein, Kagan, and Yarbrough (1977, p. 298)	Guatemala	40	5	20 (50.0%) females and 20 (50.0%) males	Auditory integration, digit memory, memory for sentences, vocabulary, matching familiar figures test score, matching familiar figures time, and embedded figures test	Children tested in Spanish by a Guatemalan psychologist.
27	Engle et al. (1977, p. 298)	Guatemala	40	7	20 (50.0%) females and 20 (50.0%) males	Auditory integration, digit memory, memory for sentences, vocabulary, matching familiar figures test score, matching familiar figures time, and embedded figures test	Children tested in Spanish by a Guatemalan psychologist.

(*table continues*)

Table 2 (*continued*)

Sample No.	Citation	Country	n	Subject age (in years)	Subject Gender	Instrument(s)	Other information
28	Engle et al. (1977, p. 299)	Guatemala	40	9	20 (50.0%) females and 20 (50.0%) males	Auditory integration, digit memory, memory for sentences, vocabulary, matching familiar figures test score, matching familiar figures time, and embedded figures test	Children tested in Spanish by a Guatemalan psychologist.
29	Engle et al. (1977, p. 299)	Guatemala	40	11	20 (50.0%) females and 20 (50.0%) males	Auditory integration, digit memory, memory for sentences, vocabulary, matching familiar figures test score, matching familiar figures time, and embedded figures test	Children tested in Spanish by a Guatemalan psychologist.
30	Fernald, Weber, Galasso, and Ratsifandrihamanana (2011, Supplementary Table 1)	Madagascar	1,232	Range: 3–6 (median is between 54 and 59 months)	645 (52.4%) females and 587 (47.6%) males	SB5 visual-spatial subtest, ESB5 working memory subtest, ESB5 matrix subtest, Leiter-R (adapted for local context), W-M memory of phrases subtest (translated and adapted into Malagasy), and day-night Stroop task (adapted)	Children selected via stratified sample. 82% of children were from rural communities.
31	Gauvain and Munroe (2009, p. 1638)	Belize, Kenya, Nepal, and American Samoa	192	Range: 5–9	96 (50.0%) females and 96 (50.0%) males	Embedded figures test, short-term memory recall of acted sequences, WPPSI block design subtest, willingness to play with toys after seeing them used, motor coordination, perspective taking test	Data were collected in the late 1970's. Each country contributed 25% of sample members. Belizean sample was a Garifuna community speaking an Arawak (native language), and Belize was a low-income country at a time. American Samoa would be too wealthy for inclusion by itself.
32	Gichuhi (1999, p. 32)	Kenya	120	Range = 12.8–16 (M = 14.3)	60 (50.0%) females and 60 (50.0%) males	12 WISC-III subtests	Examinees were candidates for the 1997 Kenya Certificate of Primary Education attending randomly selected schools in Nairobi.

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
33	Grigorenko et al. (2001, p. 374)	Kenya	86	Range = 12–15 (M = 13.4, SD = 1.03)	43 (50.0%) females and 43 (50.0%) males	Peer, teacher, and adult ratings of children's abilities, as understood in Luo culture	Sample members were members of the Luo people in a rural community. Matrix reported was not positive definite. These are likely the same individuals as in the Sternberg et al. (2001) sample.
34	Grigorenko, Ngorosho, Jukes, and Bundy (2006, p. 112)	Tanzania	194	M = 13.35, SD = 1.30	105 (54.1%) females and 89 (45.9%) males	Test of Auditory Analysis Skills (phonemic awareness); phonological memory; Test of Rapid Naming for Objects, Numbers, and Letters; digit-word naming task; reading aloud words and nonwords	Children were enrolled at schools in Bagamoyo or Kibaha Districts. All examinees tested in their native language of Kiswahili.
35	Gurven et al. (2017, p. 165)	Bolivia	461	Unknown (see other information)	461 (100.0%) females	Short-term recall, long-term recall, digit span (two tasks), spatial span, verbal fluency (three tasks), visual scan, and fluid ability	Subjects were members of the forager-horticulturalist Tsimane people in the Amazon. Age for combined male and female group is M = 33.85, SD = 18.58.
36	Gurven et al. (2017, p. 165)	Bolivia	458	Unknown (see other information)	458 (100.0%) males	Short-term recall, long-term recall, digit span (two tasks), spatial span, verbal fluency (three tasks), visual scan, and fluid ability	Subjects were members of the forager-horticulturalist Tsimane people in the Amazon. Age for combined male and female group is M = 33.85, SD = 18.58.
37	Guthrie (1963)	Philippines	314	Range = 19–22	314 (100.0%) females	50 tests detailed in Guthrie, 1963, pp. 97–98.	Correlation matrix from the American Documentation Institute Auxiliary Publications Project Document No. 7464, available from the Library of Congress.
38	Haile, Gashaw, Nigatu, and Demelash (2016, p. 147)	Ethiopia	131	Range: 6–11	Unknown	CPM and K-ABC-II triangles, word order, pattern reasoning, hand movement, Rovers, and number recall subtests	Children resided in Goba Town in southern Ethiopia. Reported age range is contradictory; may be 8–11 (see pp. 144, 145).
39	Hashmi, Tirmizi, and Shah (2010, p. 25)	Pakistan	12,120	Range = 11–12	3,853 (31.8%) females and 8,267 (68.2%) males	Spatial ability, two tests of perceptual ability, numerical ability, word fluency, and Australian matrix test	32.7% of examinees attended urban schools, while 67.3% attended rural schools.

(*table continues*)

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
40	Hashmi, Tirmizi, Shah, and Khan (2011, p. 10)	Pakistan	9,645	Range = 12–13	3,224 (33.4%) females and 6,421 (66.6%) males	Spatial ability, two tests of perceptual ability, numerical ability, word fluency, and Australian matrix test	36.9% of examinees attended urban schools, while 63.1% of examinees attended urban schools.
41	Hendrawan (2013, p. 268)	Indonesia	48	Unknown	28 (58.3%) females and 20 (41.7%) males	Letter fluency, Culture Fair Intelligence Test, WBIS similarities subtest, creativity fluency test	Subjects were healthy undergraduate students who were native speakers of Bahasa Indonesia.
42	Holding et al. (2018, p. 10)	Bangladesh	297	Range = 7–18 (*M* = 13.26, *SD* = 3.08)	167 (56.2%) females and 130 (43.8%) males	Rey-Osterrieth complex figure subtest, hand movements sequential processing test, footsteps simultaneous processing test, NOGO inhibitory control task, shift, people search selective/sustains attention, story completion planning task, Kilifi naming verbal intelligence test	Tests adapted to Bengali. Data collected in an agricultural area.
43	Holding et al. (2018, p. 10)	Ghana	166	Range = 7–18 (*M* = 12.42, *SD* = 2.81)	77 (46.4%) females and 89 (53.6%) males	Rey-Osterrieth complex figure subtest, hand movements sequential processing test, footsteps simultaneous processing test, NOGO inhibitory control task, shift, people search selective/sustains attention, story completion planning task, Kilifi naming verbal intelligence test	Examinees sampled from two schools in Kassena-Nankana District. Tests adapted to Kasim and Nankam languages. Data collected in agricultural area.
44	Holding et al. (2018, p. 10)	Tanzania	323	Range = 7–18 (*M* = 12.26, *SD* = 2.92)	164 (50.8%) females and 159 (45.0%) males	Rey-Osterrieth complex figure subtest, hand movements sequential processing test, footsteps simultaneous processing test, NOGO inhibitory control task, shift, people search selective/sustains attention, story completion planning task, Kilifi naming verbal intelligence test	Tests adapted to Kiswahili. Data collected in agricultural area, with some contributions to the economy from tourism.

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
45	Humble and Dixon (2017, p. 102)	Tanzania	1,857	Range = 7.9–19.9 (M = 11.04, SD = 1.16)	(52% female; exact count unknown)	Intelligence test, reading achievement test, mathematics achievement test, Kiswahili test	Examinees lived in Kinondoni, a poor municipality in Dar es Salaam. Students received the SPM, WAIS-II, or NNAT-2 for their intelligence test.
46	Irvine (1964, Table 4)	Zimbabwe	204	Range = 13–16 (M = 14.7)	93 (45.6%) females and 111 (54.4%) males	Practice end-of-year achievement exam, past academic achievement ("streaming"), SPM, mental alertness, compound series test, numerical ability test, Spiral Nines test	"Streaming" variable refers to the educational track that the child was in, or—if there was no ability grouping in their school—a dummy coded variable indicating whether they were in the top, middle, or bottom third of class performance. Irvine (1964, pp. 76, 628) saw this as an academic achievement variable.
47	Irvine (1964, Appendix D, Table 5)	Zimbabwe	1,615	Unknown	507 (31.4%) females and 1,108 (68.6%) males	Spiral Nines verbal and abstract tests; SPM; Normal Battery comprehension, vocabulary, and spelling subtests; and Standard 6 English (two subtests), arithmetic (two subtests), mental arithmetic, history/current affairs, geography/nature study achievement subtests	Children were randomly selected from randomly chosen schools in Mashonaland, in northern Zimbabwe. Examinees were concluding primary school at approximately age 14 or 15. Students were enrolled in urban schools (46.5%), rural boarding schools (29.4%) and rural schools (21.4%). Subjects were taking the Standard 6 examination to enter secondary school for the first time.
48	Irvine (1964, Appendix D, Table 6)	Zimbabwe	190	Unknown	Unknown	Spiral Nines verbal and abstract tests; SPM; Normal Battery comprehension, vocabulary, and spelling subtests; and Standard 6 English (2 subtests), arithmetic (2 subtests), mental arithmetic, history/current affairs, geography/nature study achievement subtests	Subjects had failed the Standard 6 examination the previous year.

(*table continues*)

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
49	Irvine (1964, Appendix F, Table 2)	Zambia	29	Unknown	Unknown	Spiral Nines verbal and abstract tests; SPM; Normal Battery comprehension, vocabulary, and spelling subtests; boxes (a 3D perceptual test); mechanical information (2 subtests); and figures (a hidden shape test)	Sample members were printers and binders.
50	Kagan et al. (1979, p. 46)	Guatemala	35	Range = 6–7	Unknown	Conservation of mass, conservation of liquid, picture span, doll span, word span, picture operations summary, doll operations summary	Sample was "approximately" half male and half female.
51	Kagan et al. (1979, p. 46)	Guatemala	35	Range = 8–9	Unknown	Conservation of mass, conservation of liquid, picture span, doll span, word span, picture operations summary, doll operations summary	Sample was "approximately" half male and half female.
52	Kagan et al. (1979, p. 47)	Guatemala	37	Range = 10–11	Unknown	Conservation of mass, conservation of liquid, picture span, doll span, word span, picture operations summary, doll operations summary	Sample was "approximately" half male and half female.
53	Kagan et al. (1979, p. 47)	Guatemala	19	Range = 12–13	Unknown	Conservation of mass, conservation of liquid, picture span, doll span, word span, picture operations summary, doll operations summary	Sample was "approximately" half male and half female.
54	Kagan et al. (1979, p. 48)	Guatemala	36	Range = 6–7	Unknown	Conservation of mass, conservation of liquid, picture span, doll span, word span, picture operations summary, doll operations summary	Sample was "approximately" half male and half female.

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
55	Kagan et al. (1979, p. 48)	Guatemala	33	Range = 8–9	Unknown	Conservation of mass, conservation of liquid, picture span, doll span, word span, picture operations summary, doll operations summary	Sample was "approximately" half male and half female.
56	Kagan et al. (1979, p. 49)	Guatemala	37	Range = 10–11	Unknown	Conservation of mass, conservation of liquid, picture span, doll span, word span, picture operations summary, doll operations summary	Sample was "approximately" half male and half female.
57	Kagan et al. (1979, p. 49)	Guatemala	34	Range = 12–13	Unknown	Conservation of mass, conservation of liquid, picture span, doll span, word span, picture operations summary, doll operations summary	Sample was "approximately" half male and half female.
58	Kearney (1966, p. 17)	Papua New Guinea	140	Unknown	140 (100.0%) males	Knox cube imitation, bead threading, passalong, form assembly, picture matching, Kim game, Gestalt continuation test	Sample members taken from various villages and agricultural plantations.
59	Lean and Clements (1981, p. 289)	Papua New Guinea	116	*M* = 19.6	2 (1.7%) females and 114 (98.3%) males	Two spatial tests, Gestalt completion test, SPM, 3D drawing test, elementary mathematics test, and mathematical processing test	Examinees were first year engineering students. 3 sample members were from the Solomon Islands, and 2 were from Samoa.
60	MacArthur, Irvine, and Brimble (1964, p. 36)	Zambia	684	*M* = 15.5 (*SD* = 1.25)	Unknown (see other information)	Spiral nines verbal and abstract scores; SPM, 1963 Junior Secondary School Leaving Examination; Normal Battery comprehension, vocabulary, spelling, and computation subtests; and Standard 6 English, arithmetic, geography, history, and science achievement test scores	Gender makeup of the sample was representative of the student population in Zambia at that time (21.5% female and 78.5% male).

(*table continues*)

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
61	MacArthur et al. (1964, p. 39)	Zambia	442	$M = 17.6$, $SD = 1.23$	Unknown	Boxes (a 3D perceptual test); mechanical information (two subtests); figures (a hidden shape test); SPM; Normal Battery mental alertness, comprehension, vocabulary, spelling, and computation subtests; 1961 Junior Secondary School Exam English, arithmetic, and special paper scores; Standard 6 Leaving Exam English, arithmetic, geography, history, and science scores; and 1963 Junior Secondary School Leaving Examination English, math, geography, and history scores	
62	MacArthur et al. (1964, p. 61)	Zambia	61	$M = 18.75$, $SD = 1.41$	Unknown	SPM; Normal Battery mental alertness, boxes (a 3D perceptual test); mechanical information (2 subtests), figures (a hidden shape test), comprehension, vocabulary, and spelling subtests; and trade efficiency ratings	Subjects were in technical school training as bricklayers and carpenters.
63	MacArthur et al. (1964, p. 62)	Zambia	57	$M = 19.67$, $SD = 1.46$	Unknown	SPM; Normal Battery mental alertness, boxes (a 3D perceptual test); mechanical information (2 subtests), figures (a hidden shape test), comprehension, vocabulary, and spelling subtests; and trade efficiency ratings	Subjects were in technical school training as bricklayers and carpenters.

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
64	MacArthur et al. (1964, p. 63)	Zambia	76	*M* = 20.58, *SD* = .91	Unknown	SPM; Normal Battery mental alertness, boxes (a 3D perceptual test); mechanical information (two subtests), figures (a hidden shape test), comprehension, vocabulary, and spelling subtests; and trade efficiency ratings	Subjects were in technical school training as bricklayers, carpenters, and mechanics.
65	Mahmood (2013, p. 98)	Pakistan	1,669	Range = 17–25	Unknown	Indigenous Intelligence Test vocabulary, arithmetic, analogy, information, comprehension, and similarity subtests	Test could be taken in English or Urdu. Examinees were students at government colleges or universities.
66	McCoy, Zuilkowski, Yoshikawa, and Fink (2017, Supplemental Table S2)	Zambia	1,623	Range = 5.58–7 (*M* = 6.22, *SD* = 4.06)	814 (50.2%) females and 809 males (49.8%)	Receptive vocabulary, letter naming, nonverbal reasoning, executive function, prosocial behavior, task orientation, and fine motor skills	All tests were new creations or adaptations for the local context of Zambia. 50.4% of the sample lived in an urban neighborhood. Examinees are a representative sample of Zambian children from all regions of the country.
67	McFie (1954, p. 39)	Uganda	30	Range = 16–20	30 (100%) females	W-B comprehension, similarities, analogies, arithmetic, picture description, picture arrangement, and block design subtests	Subjects were all student nurses. Test was translated into Luganda as needed.
68	McFie (1961, p. 237)	Uganda	26	Range = 16–19	26 (100%) males	W-B comprehension, similarities, analogies, arithmetic, picture description, picture arrangement, block design, and memory for designs subtests	Subjects were entering technical school, with 1–3 years of secondary school. All subjects tested in English.
69	Miezah (2015, p. 45)	Ghana	251	Range = 16–34	116 (46.2%) females and 135 (53.8%) males	WAIS-IV subtests	Subjects spoke English and lived in the Accra region. Minor, culturally appropriate modifications made to the information subtest.

(*table continues*)

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
70	Mourgues et al. (2016, p. 59)	Zambia	110	Range = 8–18.5 (*M* = 13.1, *SD* = 2.3)	51 (46.4%) female and 59 (53.6%) male	Phonological awareness, reading recognition, rapid automized naming, letter-digit span, pseudoword decoding, word reading, KABC-II triangles subtest, UNIT symbolic memory subtest, foreign language learning task, and Bala Bbala GraphoGame	Examinees lived in rural parts of the country.
71	Mwaura, Sylva, and Malmberg (2008, p. 245)	Kenya, Tanzania, and Uganda	423	Unknown	221 (52.2%) females and 202 (47.8%) males	British Ability Scales II block building, verbal comprehension, early number concept, and picture similarity subtests; and African Child Intelligence Test verbal meaning, exclusion, and closure subtests	378 individuals (89.4%) were between 3 and 5 years old. 34 (8.0%) were older than 5.
72	O'Donnell et al. (2012, Table 3)	Cambodia, Ethiopia, India, Kenya, and Tanzania	1,206	Range = 6–12	573 (47.5%) females and 633 (52.5%) males	KABC-II hand movements, triangles, and pattern reasoning subtests; adapted/abbreviated version of California Verbal Learning Test-Children's Version	All children were abandoned or orphaned
73	Orbell (1981, p. 12)	Zimbabwe	170	Unknown	49 (40.5%) females and 121 (71.2%) males	SPM, mental alertness, English comprehension, English vocabulary, syllogistic-type reasoning, Piagetian tasks	Subjects were all native Shona speakers. No subject was younger than 17, and most subjects were probably 18.
74	Ord (1970, p. 46)	Papua New Guinea	100	Unknown	Unknown	Cube imitation, bead threading, passalong, form assembly, and observation	Subjects were all military members.
75	Panza Lombardo (2016, p. 55)	Paraguay	75	Range = 18–37, *M* = 23.08, *SD* = 3.7	45 (60.0%) females and 30 (40.0%) males	SPM, CFIT-III, WAIS digit span subtest, Reading Span Test	Subjects were English-speaking university students. Reading Span Test was adapted into Spanish.

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
76	Peters, Baker, Dieckmann, Leon, and Collins (2010, p. 1372)	Ghana	181	Range = 30–75, *M* = 43.9, *SD* = 10.58	130 (71.8%) females and 51 (28.2%) males	WAIS backward digit span subtest, Delis-Kaplan executive function system tower test, WJ-III calculation subtest, stickman test, HIV/AIDS knowledge, and protective behaviors	Examinees lived in four rural villages, and 48.9% farmed as their primary occupation. All tests administered in the Twi language. Stickman test was developed for this culture.
77	Pitchford and Outhwaite (2016, p. 9)	Malawi	283	Range = 6.08–13.42 (*M* = 8.10, *SD* = 1.26)	139 (49.1%) females and 144 (50.9%) males	Manual processing speed, manual coordination, short-term memory, visual attention, backward spatial memory span, and 2D block design task	Participants were children attending an urban state school in Lilongwe. Children were tested via an electronic system.
78	Pongcharoen et al. (2011)	Thailand	560	*M* = 9.28, *SD* = .33	276 (49.3%) females and 284 (50.7%) males	CPM, WISC-III picture completion, information, coding, similarity, picture arrangement, and arithmetic subtests	All children were tested at 9 years. 75.2% of children were receiving iron and/or zinc supplements as part of an intervention. Correlation matrix obtained via personal communication (T. Pongcharoen, February 7, 2018).
79	Rasheed et al. (2017, p. 7)	Pakistan	1,273	4	586 (46.0% females and 687 (54.0%) males	WPPSI-III subtests	All children were tested at 4 years in Naushero Feroze, a rural district. Some subtests adapted to fit the cultural and linguistic context
80	Rehna and Hanif (2017, p. 25)	Pakistan	300	Range = 12–16 (*M* = 14.83, *SD* = 1.16)	151 (50.3%) females and 149 (49.7%) males	Word memories, verbal reasoning, numerical ability, general knowledge, and verbal ability	All tests were subtests of the Sajjad Verbal Intelligence Test Urdu.
81	Reyes et al. (2010, Table 2)	Philippines	742	2	340 (45.8%) females and 402 (54.2%) males	Griffiths Scales of Mental Development locomotor, personal-social, eye-hand coordination, hearing and language, and performance development subscales	All children were tested at 4 years. Subjects lived in the province of Bulacan.

(*table continues*)

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
82	Ruffieux et al. (2010)	Cameroon	134	Range = 4–20 (*M* = 11.48, *SD* = 4.24)	74 (55.2%) females and 60 (44.8%) males	Purdue Pegboard Test (five variables), California Verbal Learning Test for children (six variables), Color Trails 1 and 2, K-ABC hand movements subtest, digit span, backward digit span, WISC-IV coding subtest, WISC-IV letter-number sequencing subtest, Verbal Semantic Fluency Test, Bell Cancellation Task, and the Continuous Performance Test II (five variables)	Subjects attended public schools in Yaoundé. Testing occurred in French, which was a second language for most children. Correlation matrix obtained via personal communication (N. Ruffieux, personal communication, February 19, 2018). Sample size is 134 in this report (instead of 125 in the article) because of the inclusion here of data from nine children ages 4–5.
83	Sen, Jensen, Sen, and Arora (1983, p. 148)	India	45	Range = 18–50 (*M* = 27.92, *SD* = 7.33)	45 (100.0%) males	SPM, forward digit span, backward digit span, and six reaction time tasks	Subjects were employees of Delhi University.
84	Shrestha (1994, p. 50, Table 5a)	Malawi	72	Range = 6–8 (M = 7.1, SD = .7)	Unknown	Verbal fluency (two tests), exclusion task, quantity task, verbal meaning, visual memory, and closure	All children were N'goni individuals who lived five villages in Ntcheu District.
85	Songy (2007, p. 67)	Papua New Guinea	135	*M* = 24.04, *SD* = 3.17	Unknown (presumably 100% male)	Grade-point average; CFIT-II series, classification, matrices, and conditions subtests; CFIT-III series, classification, matrices, and conditions subtests; Michigan Test of English Language Proficiency vocabulary, grammar, and comprehension subtests	Subjects were all English-speaking Catholic seminarians and had completed secondary education.
86	Stemler et al. (2009, p. 174)	Zambia	100	Unknown	Unknown	ZAT mathematics achievement, reading recognition, reading comprehension, and pseudoword decoding subtests	Students tested in English. This is part of a larger matrix of cognitive variables, but the sample description is not clear enough to determine whether other portions of the matrix refer to the same sample.

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
87	Stemler et al. (2009, p. 174)	Zambia	497	Unknown	Unknown	ZAT mathematics achievement, reading recognition, reading comprehension, and pseudoword decoding subtests	Students tested in Nyanja. This is part of a larger matrix of cognitive variables, but the sample description is not clear enough to determine whether other portions of the matrix refer to the same sample.
88	Stemler et al. (2009, p. 177)	Zambia	115	Unknown	Unknown	Mathematics achievement (two measures), overall English achievement, reading recognition, pseudoword decoding, and reading comprehension	Examinees were enrolled in Grade 5.
89	Sternäng, Lövdén, Kabir, Hamadani, and Wahlin (2016)	Bangladesh	452	Range = 60–92 (M = 69.3, SD = 6.8)	55% female and 45% male[b]	Recall (two tests), recognition (two tests), verbal fluency (two tests), and processing speed (two tests) tasks	Correlation matrix obtained via personal communication (O. Sternäng, personal communication, January 19, 2018).
90	Sternberg et al. (2001, p. 411)	Kenya	85	Range = 12–15	42 (49.4%) females and 43 (50.6%) males	Tacit knowledge for natural herbal medicines, Dholuo vocabulary scale, English Mill Hill vocabulary scale, and CPM	Individuals were members of the Luo people in a rural community. These are likely the same individuals as in the Grigorenko et al. (2001) sample.
91	Sternberg et al. (2002, p. 154)	Tanzania	458	Range = 11–13	257 (56.1%) females and 201 (43.9%) males	Syllogisms, Wisconsin Card Sorting Test (two variables), and Twenty Questions task	Pre-test data only included in the matrix. Data combine experimental and control samples. Children lived in a rural area.
92	Sukhatunga et al. (2002, p. 352)	Thailand	1,525	Range = 6–16	Unknown	WASI vocabulary, similarities, block design, and matrix reasoning subtests	Test subscales were translated and adapted into Thai.
93	Sukhatunga et al. (2002, p. 352)	Thailand	792	Range = 17–70	Unknown	WASI vocabulary, similarities, block design, matrix reasoning, information, and picture completion subtests	Test subscales were translated and adapted into Thai.
94	Tan, Reich, Hart, Thuma, and Grigorenko (2014, p. 276)	Zambia	114	M = 12.094, SD = 2.34	57 (50%) females and 57 (50%) males	Phonological awareness, UNIT symbolic memory subtest, and ZAT reading recognition, reading comprehension, and mathematics achievement subtests	Examinees lived in a rural area.

(*table continues*)

Table 2 (*continued*)

Sample No.	Citation	Country	*n*	Subject age (in years)	Subject Gender	Instrument(s)	Other information
95	van de Vijver (2002)	Zambia	376	Unknown	Unknown	Inductive reasoning (two tasks), rule classification, rule generating, and rule testing	Correlation matrix obtained via personal communication (F. van de Vijver, personal communication, January 17, 2018).
96	van de Vijver (2002)	Zambia	328	Unknown	Unknown	Inductive reasoning (two tasks), rule classification, rule generating, and rule testing	Correlation matrix obtained via personal communication (F. van de Vijver, personal communication, January 17, 2018).
97	van den Briel et al. (2000, p. 1182)	Benin	196	Range = 7–11 (M = 8.9, SD = 1.2)	29 (14.8%) females and 167 (85.2%) males	WPPSI block design subtest; African Child Intelligence Test closure, concentration, exclusion, fluency, and mazes subtests; K-ABC hand movements subtest, and CPM	Children lived in the district of Basila (province of Atacora).
98	Veena (2014, p. 56)	India	542	Range = 9–10 (M = 9.5)	281 (48.1%) females and 261 (51.9%) males	K-ABC-II Atlantis, word order, and pattern reasoning subtests; verbal fluency; Kohs block design test; and WISC-III coding subtest	Tests administered in the Kannada language.
99	Warburton (1951, p. 128)	Nepal	439	Unknown	400 (100.0%) males	Screwboard (manual dexterity test), formboard, nailboard (tracing task), block design, color series (pattern completion), cube construction, and dice (simplified version of SPM)	All men were soldiers. Sample size in article is inconsistent. Maximum sample size (taken from Table 7) is used here.

Note. CFIT-II = Cattell Culture Free Intelligence Test, Second Edition; CFIT-III = Cattell Culture Free Intelligence Test, Third Edition; CPM = Raven's Colored Progressive Matrices; ESB5 = Stanford-Binet Intelligence Scales for Early Childhood, Fifth Edition; K-ABC = Kaufman Assessment Battery for Children; Leiter-R = Leiter International Performance Scale—Revised; NNAT-2 = Naglieri Nonverbal Ability Test; SB5 = Stanford-Binet Intelligence Scales, Fifth Edition; SPM = Raven's Standard Progressive Matrices; UNIT = Universal Nonverbal Intelligence Test; WAIS = Wechsler Adult Intelligence Scale; WAIS-II = Wechsler Adult Intelligence Scale—Second Edition; WAIS-IV = Wechsler Adult Intelligence Scale—Fourth Edition; W-B = Wechsler-Bellevue Intelligence Scale; WISC-R = Wechsler Intelligence Scale for Children—Revised; WISC-III = Wechsler Intelligence Scale for Children—Third Edition; WJ-III = Woodcock-Johnson Tests of Cognitive Ability, Third Edition; W-M = Woodcock-Muñoz Language Survey; WPPSI = Wechsler Preschool and Primary Scale of Intelligence; WPPSI-III = Wechsler Preschool and Primary Scale of Intelligence—Third Edition; ZAT = Zambian Achievement Test.

[a] Berry et al. (1986) reported separate matrices of Bangandu (p. 171), Biaka (p. 171), and Gbanu individuals. The Bangandu and Biaka matrices were not positive definite, due to the use of pairwise deletion. [b] Exact counts of males and females unknown.

Table 3
Dataset Information and Factor Analysis Results

Sample No.	Modified Guttman No. of factors	MAP No. of factors	% of variance explained in first unrotated factor	Schmid-Leiman factor analysis results
1	2	1	27.452%	A general factor accounted for 30.8% extracted variance. Two first-order factors accounted for 69.2% of extracted variance.
2	1	1	46.174%	—
3	1	1	46.399%	—
4	2	1	38.606%	A general factor accounted for 14.6% extracted variance. Two first-order factors accounted for 85.4% of extracted variance.
5	1	1	47.362%	—
6	2	1	18.348%	A general factor accounted for 17.0% extracted variance. Two first-order factors accounted for 83.0% of extracted variance.
7	2	1	24.735%	A general factor accounted for 43.2% extracted variance. Two first-order factors accounted for 56.8% of extracted variance.
8	1	1	39.216%	—
9	3	1	28.676%	A general factor accounted for 19.8% extracted variance. Three first-order factors accounted for 80.2% of extracted variance.
10	1	1	52.252%	—
11[a]	2	1	44.658%	A general factor accounted for 28.0% extracted variance. Two first-order factors accounted for 72.0% of extracted variance.
12[b]	1	1	49.260%	—
13	1	1	53.565%	—
14	—	—	—	—
15	4	1	39.459%	A general factor accounted for 63.2% extracted variance. Four first-order factors accounted for 36.8% of extracted variance.
16	1	2	63.479%	A general factor accounted for 71.1% extracted variance. Two first-order factors accounted for 28.9% of extracted variance.
17	1	1	49.451%	—
18	1	1	44.568%	—
19	1	1	44.375%	—
20	1	1	39.677%	—
21	2	3	38.653%	A general factor accounted for 55.2% extracted variance. Two first-order factors accounted for 44.8% of extracted variance. A general factor accounted for 54.5% extracted variance. Three first-order factors accounted for 45.5% of extracted variance.
22	1	1	56.864%	—
23	1	1	56.925%	—
24	1	1	79.592%	—
25	1	1	41.872%	—
26	1	1	44.600%	—
27	1	1	49.270%	—
28	1	1	31.633%	—
29	1	1	35.320%	—
30	1	1	41.685%	—
31	1	1	48.173%	—
32[c]	1	2	46.801%	A general factor accounted for 67.7% extracted variance. Two first-order factors accounted for 32.3% of extracted variance.
33	—	—	—	—
34	3	2	31.316%	A general factor accounted for 49.9% extracted variance. Two first-order factors accounted for 50.1% of extracted variance. Two general factors accounted for 67.3% extracted variance. Three first-order factors accounted for 32.7% of extracted variance.
35	2	2	35.270%	Factor extraction not possible for a 1, 2, or 3-factor solution.
36	2	2	34.880%	Factor extraction not possible for a 1, 2, or 3-factor solution.
37	10	5	18.807%	A general factor accounted for 42.6% extracted variance. Five first-order factors accounted for 57.4% of extracted variance. A general factor accounted for 34.6% extracted variance. Ten first-order factors accounted for 65.4% of extracted variance.
38	1	1	48.933%	—
39	3	2	45.140%	A general factor accounted for 34.1% extracted variance. Two first-order factors accounted for 65.9% of extracted variance. Factor extraction not possible for a 3-factor solution.
40	2	1	37.893%	A general factor accounted for 13.0% extracted variance. Two first-order factors accounted for 87.0% of extracted variance.
41	1	1	59.417%	—
42	2	2	41.769%	A general factor accounted for 50.4% extracted variance. Two first-order factors accounted for 49.6% of extracted variance.

(*table continues*)

Table 3 (*continued*)

Sample No.	Modified Guttman No. of factors	MAP No. of factors	% of variance explained in first unrotated factor	Schmid-Leiman factor analysis results
43	1	1	33.855%	—
44	1	1	38.754%	—
45	1	1	49.494%	—
46	1	1	55.724%	—
47	2	1	42.423%	A general factor accounted for 73.4% extracted variance. Two first-order factors accounted for 26.6% of extracted variance.
48	1	1	42.374%	—
49	1	2	58.359%	A general factor accounted for 68.9% extracted variance. Two first-order factors accounted for 31.1% of extracted variance.
50	1	1	41.860%	—
51	1	1	48.549%	—
52	1	1	49.280%	—
53	1	1	49.192%	—
54	1	1	37.930%	—
55	1	2	39.491%	A general factor accounted for 23.2% extracted variance. Two first-order factors accounted for 76.8% of extracted variance.
56	1	1	34.814%	—
57	1	1	43.922%	—
58	1	1	47.776%	—
59[d]	1	1	36.349%	—
60	2	1	42.792%	A general factor accounted for 69.5% extracted variance. Two first-order factors accounted for 30.5% of extracted variance.
61	5	3	25.473%	A general factor accounted for 46.1% extracted variance. Three first-order factors accounted for 53.9% of extracted variance. For a five-factor model, a general factor accounted for 42.9% of variance. Five first-order factors accounted for 57.1% of variance.
62	1	1	35.557%	—
63	2	2	33.450%	A general factor accounted for 39.7% extracted variance. Two first-order factors accounted for 60.3% of extracted variance.
64	1	1	40.067%	—
65[e]	1	1	45.598%	—
66	1	1	35.577%	—
67	1	1	47.705%	—
68	1	1	30.275%	—
69	1	1	49.562%	—
70	1	1	55.107%	—
71	1	1	50.020%	—
72	1	1	56.096%	—
73	1	1	35.240%	—
74	1	1	46.777%	—
75	1	1	84.186%	—
76	1	1	43.910%	—
77	1	1	26.079%	—
78	1	1	42.195%	—
79	1	1	39.824%	—
80	1	1	86.301%	—
81	1	1	55.476%	—
82	3	3	46.998%	A general factor accounted for 69.9% extracted variance. Three first-order factors accounted for 30.1% of extracted variance.
83	1	1	41.346%	—
84	1	1	32.774%	—
85	2	2	26.745%	A general factor accounted for 28.1% extracted variance. Two first-order factors accounted for 71.9% of extracted variance.
86	1	1	71.980%	—
87	1	1	73.688%	—
88	1	1	75.209%	—
89	2	2	45.986%	A general factor accounted for 59.6% extracted variance. Two first-order factors accounted for 40.4% of extracted variance.
90[f]	1	1	48.677%	—
91	1	1	34.094%	—
92	1	1	55.501%	—
93	1	1	65.699%	—
94	1	1	61.922%	—
95	1	1	63.646%	—
96	1	1	62.293%	—

Table 3 (*continued*)

Sample No.	Modified Guttman No. of factors	MAP No. of factors	% of variance explained in first unrotated factor	Schmid-Leiman factor analysis results
97	1	1	35.986%	—
98[g]	1	1	41.940%	—
99	1	1	41.888%	—

[a] Berry et al. (1986) reported separate matrices of Bangandu (p. 171), Biaka (p. 171), and Gbanu individuals. The Bangandu and Biaka matrices were not positive definite, due to the use of pairwise deletion. The Gbanu matrix (p. 173) contained four variables and resulted in one factor (regardless of factor number selection method) that resulted in 41.083% of observed variable variance. [b] Bhatia (1955, p. 50) presents an alternate correlation matrix that has age partialled out. The results indicate one factor (regardless of factor number selection method) that explains 44.511% of observed variable variance. [c] Gichuhi (1999, p. 32) also reported a correlation matrix based on standardized WISC-III subtest scores that were transformed using Fisher's *z* transformation and the American WISC-III standardization data. The results indicate one factor (regardless of factor number selection method) that explains 46.133% of observed variable variance. [d] Lean and Clements (1981, p. 290) also reported an alternative correlation matrix which included two tests of mathematics and combined the two spatial ability test scores. The results indicate one factor (regardless of factor number selection method) that explains 33.030% of observed variable variance. [e] Mahmood (2013, p. 99) also includes the correlation between subtest scores and college grades. The results indicate one factor (regardless of factor number selection method) that explains 40.941% of observed variable variance. [f] Sternberg et al. (2001, p. 412) also reported an alternate correlation matrix with age and socioeconomic status partialled out. The results indicate one factor (regardless of factor number selection method) that explains 43.715% of observed variable variance. [g] Veena (2014, p. 56) also reported a correlation matrix of the same variables measured when the same subjects were in adolescents (age 13–14). The results indicate one factor (regardless of factor number selection method) that explains 41.746% of observed variable variance.

first-order factors when using the modified Guttman rule. When the modified Guttman rule's three factors are analyzed using a second-order EFA, two factors emerged. However, when MAP's two factors are analyzed using a second-order EFA, only one factor emerged. Thus, even this dataset cannot consistently produce multiple higher-order factors.

Table 3 also displays the percentage of variance accounted for by the first extracted factor in the initial EFA of each dataset. The average first factor explained 45.9% of variance ($SD = 12.9\%$), with a median of 44.4%. The weakest first factor accounted for 18.3% of variance (Attallah, Ahmed, & Meisenberg, 2014), while the strongest first factor accounted for 86.3% of observed variable variance (Rehna & Hanif, 2017).

Thirty-one second-order EFAs of 26 data sets were attempted after more than one first-order factor was obtained. Five data sets (Dasen, 1984; Grigorenko et al., 2006; Guthrie, 1963; Hashmi, Tirmizi, & Shah, 2010; MacArthur, Irvine, & Brimble, 1964) were given multiple second-order EFAs because there were differing numbers of first-order factors retained when using the modified Guttman and MAP rules. Of these 31 second-order EFAs, factor extraction was not possible in three of them (both Gurven et al., 2017, samples and the three-factor solution for Hashmi et al., 2010), and—as already stated—one produced two second-order factors (Grigorenko et al., 2006, three first-order factor solutions). Of the remaining second-order EFAs, a general factor accounted for an average 44.9% of extracted variance ($SD = 19.1\%$).

Discussion

General Discussion of Results

We conducted this study to create a strong test of the theory that general cognitive ability is a cross-cultural trait by searching for *g* in human populations where *g* would be the least likely to be present or would be weakest. The results of this study are remarkably similar to results of EFA studies of Western samples, which show that *g* accounts for approximately half of the variance among a set of cognitive variables (e.g., Canivez & Watkins, 2010). In our study, the first extracted factor in 97 EFAs of data sets from non-Western, nonindustrialized countries was 45.9%.

Moreover, 73.2% of the data sets unambiguously produced a single factor, regardless of the method used to select the number of factors in the EFA. Of the remaining data sets, almost every one in which a second-order EFA was possible produced a single general factor. The only exceptions were from Grigorenko, Ngorosho, Jukes, and Bundy (2006) and Gurven et al. (2017). The Grigorenko et al. (2006) dataset produced two general factors only if one sees the modified Guttman method of selecting the number of first-order factors as being a more realistic solution than MAP. Given the modified Guttman rule's penchant for overfactoring and the generally accurate results from MAP in simulation studies (Warne & Larsen, 2014), it seems more likely that even the Grigorenko et al. (2006) dataset has two first-order factors and one general factor that accounts for 49.9% of extracted variance. The Gurven et al. (2017) samples both produced two factors in an initial EFA, but the factor extraction process failed for the second-order EFAs for both samples. The inability to test whether the two initial factors could form a general factor makes the Gurven et al. (2017) data ambiguous in regards to the evidence of the presence of *g* in its Bolivian samples.

Although we did not preregister any exact predictions for our study, we are astonished at the uniformity of these results. We expected before this study began that many samples would produce *g*, but that there would have been enough samples for us to conduct a *post hoc* exploratory analysis to investigate why some samples were more likely to produce *g* than others. With only three samples that did not produce *g*, we were unable to undertake our plans for exploratory results because *g* appeared too consistently in the data.

Thus, Spearman's *g* appeared in at least 94 of the 97 data sets (97.0%) from 31 countries that we investigated, and the remaining three samples produced ambiguous results. *Because these data sets originated in cultures and countries where g would be least likely to appear if it were a cultural artifact, we conclude that general cognitive ability is likely a universal human trait.* The characteristics of the original studies that reported these data support this conclusion. For example, some of these data sets were collected by individuals who are skeptical of the existence or primacy of *g* in general or in non-Western cultures (e.g., Hashmi et al., 2010; Hashmi, Tirmizi, Shah, & Khan, 2011; O'Donnell et al., 2012; Pitchford & Outhwaite, 2016; Stemler et al., 2009; Sternberg et al., 2001, 2002). One would think that these investigators would be most likely to include variables in their data sets that would form an additional factor. Yet, with only three ambiguous exceptions (Grigorenko et al., 2006; Gurven et al., 2017), these researchers' data still produced *g*. Additionally, many of these data sets were collected with no intention of searching for *g* (e.g., Bangirana et al., 2015; Berry et al., 1986; Engle, Klein, Kagan, & Yarbrough, 1977; Kagan et al., 1979; McCoy, Zuilkowski, Yoshikawa, & Fink, 2017; Mourgues et al., 2016; Ord, 1970; Rehna & Hanif, 2017; Reyes et al., 2010; Tan, Reich, Hart, Thuma, & Grigorenko, 2014). And yet a general factor still developed anyway. It is important to recognize, though, that the *g* factor explained more observed variable variance in some samples than in others.

For those who wish to equate *g* with a Western view of "intelligence," this study presents several problems for the argument that Western views of intelligence are too narrow. First, in our search, we discovered many examples of non-Western psychologists using Western intelligence tests with little adaptation and without expressing concern about the tests' overly narrow measurement techniques. Theorists who argue that the Western perspective of intelligence is too culturally narrow must explain why these authors use Western (or Western-style) intelligence tests and why these tests have found widespread acceptance in the countries we investigated (Oakland, Douglas, & Kane, 2016). Another difficulty for the argument that Western views of intelligence are too narrow is the fact that tests developed in these nonindustrialized, non-Western cultures positively correlate with Western intelligence tests (Mahmood, 2013; van den Briel et al., 2000). This implies that these indigenous instruments are also *g*-loaded to some extent, which would support Spearman's (1927) belief in the indifference of the indicator.

One final issue bears mention. Two peer reviewers raised the possibility that developmental differences across age groups could be a confounding variable because a *g* factor may be weaker in children than adults. To investigate this possibility, we conducted two *post hoc* nonpreregistered analyses. First, we found the correlation between the age of the sample (either its mean or the midpoint of the sample's age range) and the variance explained by the first factor in the dataset was $r = .127$ ($r^2 = .016$, $n = 84$, $p = .256$). Because a more discrete developmental change in the presence of strength of a *g* factor was plausible, we also divided the data sets five age groups: <7 years (10 samples), 7–12.99 years (34 samples), 13–17.99 years (12 samples), 18–40.99 years (21 samples), and ≥41 years (five samples). All of these age groups had a mean first factor that had a similar strength (between 41.79% and 49.63%), and the null hypothesis that all age groups had statistically equal means could not be rejected ($p = .654$, $\eta^2 = .031$) These analyses indicate that there was no statistical relationship between sample age and the strength of the *g* factor in a dataset.

Methodological Discussion

A skeptic of *g* could postulate that our results are a statistical artifact of the decisions we used to conduct a factor analysis. Some data sets in our study had been subjected to EFA in the past, and the results often differed from ours (Attallah et al., 2014; Bulatao & Reyes-Juan, 1968; Church, Katigbak, & Almario-Velazco, 1985; Conant et al., 1999; Dasen, 1984; Dawson, 1967b; Elwan, 1996; Guthrie, 1963; Humble & Dixon, 2017; Irvine, 1964; Kearney, 1966; Lean & Clements, 1981; McFie, 1961; Miezah, 2015; Orbell, 1981; Rasheed et al., 2017; Ruffieux et al., 2010; Sen, Jensen, Sen, & Arora, 1983; Sukhatunga et al., 2002; van en Briel et al., 2000; Warburton, 1951). In response, we wish to emphasize that we chose procedures *a priori* that are modern methods accepted among experts in factor analysis (e.g., Fabrigar et al., 1999; Larsen & Warne, 2010; Thompson, 2004; Warne & Larsen, 2014). The use of promax rotation, for example, might be seen as an attempt to favor correlated first-order factors—which are mathematically much more likely to produce a second-order *g* than orthogonal factors. However, promax rotation does not force factors to be correlated, and indeed uncorrelated factors are possible after a promax rotation. Therefore, the use of promax rotation permitted a variety of potential factor solutions—including uncorrelated factors—and permitted the strong test of *g* theory that we desired.

Another potential source of criticism would be our methods of retaining the number of factors in a dataset. The original Guttman (1954) rule of retaining all factors with an eigenvalue of 1.0 or greater is the most common method used in the social sciences, probably because it is the default method on many popular statistical analysis packages (Fabrigar et al., 1999). However, the method can greatly overfactor, especially when a dataset has a large number of variables, the sample size is large, and when factor loadings are weak (Warne & Larsen, 2014). These circumstances are commonly found in cognitive data sets, which are frequently plagued by overfactoring (Frazier & Youngstrom, 2007). This is why we chose to use more conservative and accurate methods of retaining the number of factors (Warne & Larsen, 2014). The use of MAP is especially justified by its strong performance in simulation studies and its tendency to rarely overfactor. MAP is insensitive to sample size, the correlation among observed variables, factor loading strength, and the number of observed variables (Warne & Larsen, 2014), all of which varied greatly among the 97 analyzable data sets.

Indeed, it is because of our use of modern methods of factor selection and rotation that we believe that prior researchers have never noticed *g* as a ubiquitous property of cognitive data in non-Western groups. Many prior researchers used varimax rotation and the original Guttman rule, likely because these methods mathematically and computationally were easier in the days before inexpensive personal computers or because both are the default method in popular statistics packages today. (Additionally, the older data sets predate the invention of promax rotation and/or MAP). But both of these methods obscure the presence of *g*. As an extreme example, Guthrie's (1963) data consist of 50 observed variables (the most of any dataset in our study) that produced 22 factors when he subjected them to these procedures. Some of Guthrie's (1963) factors were weak, uninterpretable, or defined by just one or two variables. In our analyses we found five (using

MAP) or 10 (using the modified Guttman rule) first-order factors; when subjected to the second-order EFA, the data clearly produced a single factor with an obvious interpretation: *g*.

The results of this study are highly unlikely to be a measurement artifact because the original researchers used a wide variety of instruments to measure cognitive skills in examinees. While some of these instruments were adaptations of Western intelligence tests (e.g., Abdelhamid, Gómez-Benito, Abdeltawwab, Bakr, & Kazem, 2017), some samples included variables that were based on Piagetian tasks (e.g., Dasen, 1984; Kagan et al., 1979; Orbell, 1981). Other samples included variables that were created specifically for the examinees' culture (e.g., Mahmood, 2013; Stemler et al., 2009; Sternberg et al., 2001; van den Briel et al., 2000) or tasks that did not resemble Western intelligence test subtests (Bangirana et al., 2015; Berry et al., 1986; Gauvain & Munroe, 2009). There were also several samples that included measures of academic achievement in their data sets (e.g., Bulatao & Reyes-Juan, 1968; Guthrie, 1963; Irvine, 1964). The fact that *g* emerged from such a diverse array of measurements supports Spearman's (1927) belief in the "indifference of the indicator" and shows that any cognitive task will correlate with *g* to some degree.

Other readers may object to our use of EFA at all, arguing that a truly strong test of *g* theory would be to create a confirmatory factor analysis (CFA) model in which all scores load onto a general factor. However, we considered and rejected this approach because CFA only tests the model(s) at hand and cannot generate new models from a dataset (Thompson, 2004). In this study, EFA procedures did not "know" that we were adherents to *g* theory when producing the results. Rather, "EFA methods . . . are designed to 'let the data speak for themselves,' that is, to let the structure of the data suggest the most probable factor-analytic model" (Carroll, 1993, p. 82). Thus, if a multifactor model of cognitive abilities were more probable in a dataset than a single *g* factor, then EFA would be more likely to identify it than a CFA would. The fact that these EFAs so consistently produced *g* in their data is actually a stronger test of *g* than a set of CFAs would have been because EFA was more likely to produce a model that disproved *g* than a CFA would. CFA is also problematic in requiring the analyst to generate a plausible statistical model—a fact that Carroll (1993, p. 82) recognized when he wrote:

> It might be argued that I should have used CFA. . . . But in view of wide variability in the quality of the analyses applied in published studies, I could not be certain about what kind of hypotheses ought to be tested on this basis. (Carroll, 1993, p. 82)

We agree with Carroll on this point. CFA also requires exactly specifying the appropriate model(s) to be tested. While this is a positive aspect of CFA in most situations, it was a distinct disadvantage when we were merely trying to establish whether *g* was present in a dataset that may not have been collected for that purpose. This is because most authors usually did not report a plausible theoretical model for the structure of their observed variables, and there was often insufficient information for us to create our own plausible non-*g* models that could be compared with a theory of the existence of Spearman's *g* in the data.[3] Indeed, some researchers did not collect their data with any model of intelligence in mind at all (e.g., McCoy et al., 2017). By having EFA to generate a model for us, we allowed plausible competing models to emerge from each dataset and examined them afterward to see if they supported our theory of the existence of Spearman's *g* in non-Western cultures. Another problem with CFA's requirement of prespecified models is that some theories of cognitive abilities include *g* as part of a larger theoretical structure of human cognition (e.g., Canivez, 2016; Carroll, 1993). How the non-*g* parts of a model might relate to *g* and to the observed variables is rarely clear.

Another advantage to EFA over CFA is that the former uses data to generate a new model atheoretically, and the subjective decisions (e.g., factor rotation method, second-order procedures, standards used to judge the number of factors) in an EFA are easily preregistered, whereas the subjective decisions in a CFA (e.g., when to use modification indices, how to arrange variables into factors, the number of non-*g* factors to include in a model) often cannot be realistically preregistered—or even anticipated before knowing which variables were collected—in secondary data analysis if the data were not collected in a theoretically coherent fashion (as was often the case for our data sets). By preregistering the subjective decisions in an EFA, we could ensure that subjective decisions could not bias our results into supporting our preferred view of cognitive abilities.

Finally, we want to remind readers that our dataset search and analysis procedures were preregistered and time stamped at the very beginning of the study before we engaged in any search procedures or analyses. This greatly reduces the chance for us to reverse engineer our methods to ensure that they would produce the results we wanted to obtain. Still, deviations from our preregistration occurred. When we deviated from the preregistration protocol, we stated so explicitly in this article, along with our justification for the deviation. Additionally, some unforeseen circumstances presented themselves as we conducted this study. When these circumstances required subjective decisions after we had found the data, we erred on the side of decisions that would maximize the chances that the study would be a strong test of *g* theory. Again, we have been transparent about all of these unforeseen circumstances and the decisions we made in response to them.

Limitations

Readers should be aware of limitations in our study. The first important limitation is the difficulty in finding data sets from many countries on the list in Table 1. This happened for a variety of reasons. For example, we were limited to studies published in English and Spanish (the languages that we are fluent in), but we also found promising references to works in Arabic, French, and Chinese that could possibly contain relevant data. Another hurdle was the lack of published correlation matrices in many reports and

[3] A typical example is the Bau and Dyck (1992) study, the researchers collected nine variables that were a mix of academic achievement; verbal, mathematical, and general reasoning; and speed and accuracy tasks. Beyond reporting a correlation matrix, the authors never specified the relationships of these variables to one another. How these nine variables could group into multiple factors is not at all clear, and any model we could create would be arbitrary. It is also not clear *a priori* whether variable error terms need to be correlated because of shared method or some subtests being administered in an examinee's non-native language (in this case, English). Thus, the characteristics of a plausible CFA model are uncertain, and creating such a model is fraught with many more subjective decisions than in an EFA.

the reluctance of some authors—including some of the most prominent scholars in the areas of international cognitive and educational assessment—to share their data with us. Finally, many relevant data sets were published in sources that are not indexed by traditional databases. Google Scholar was an effective method of finding data sets that were published in regional or national journals, book chapters, predatory journals, and dissertations. But there are likely many more data sets that we were unable to find because no database index is comprehensive.

Another limitation is that we did not include studies of individuals from the countries in Table 1 who had immigrated to other nations or the children of such immigrants (e.g., Lynn & Owen, 1994). Such studies could provide useful information, but we eliminated them *a priori* from our research because we worried that the acculturation process could have made *g* develop in cognitive data from these individuals. Another potentially interesting population to study could be indigenous groups within countries that are not in Table 1, such as South Africa, Brazil, Mexico, or China. Each of these countries is not classified by the United Nations as a least developed country or a low middle income nation, but rural areas of these nations may have populations that have experienced few of the cultural or economic influences of Westernization. Also, individuals from the nations of the Arab world (where economic prosperity derived from petroleum deposits) and the Caribbean (where strong tourism-based economies disqualified many countries from our study) may be valuable to investigate.

Although we believe that this study establishes the presence of *g* in data from these non-Western cultures, this study says nothing about the relative level of general cognitive ability in various societies, nor can it be used to make cross-cultural comparisons. For this purpose, one must establish measurement invariance of a test across different cultural groups (e.g., Holding et al., 2018) to ensure that test items and tasks function in a similar way for each group. The existence of *g* in all of these cultures shows that—contrary to the opinions of cultural relativists (e.g., Berry, 1986)—such comparisons may be possible, though the practical hurdles to creating tests for dissimilar cultures show that establishing measurement invariance across groups may be very difficult to overcome (Church & Katigbak, 1987). Even popular tests that are supposedly usable in many cultures—such as the draw-a-person test or the Raven's matrices—sometimes function poorly when administered to non-Western examinees (Dutton et al., 2018; Kathuria & Serpell, 1998; Wicherts et al., 2010a), and considerable work may be necessary to measure the cultural manifestations and indicators of general cognitive ability in some groups. Investigators who wish to make cross-cultural comparisons must demonstrate that (a) their tests are appropriate for examinees' cultures, (b) examinees understand the instructions and how to respond to test items, and (c) the tests demonstrate measurement invariance across groups. None of these requirements can be assumed without evidence.

We also want to make it clear that even though *g* likely exists in many human populations, this study does not say how *g* relates to other cognitive abilities or whether the structure of abilities is the same across cultures. An investigation using multigroup confirmatory factor analysis of the same variables would be capable of determining whether the relationships among *g*, non-*g* abilities, and observed variables is the same across cultural groups.

A final limitation of our work is that there is no guarantee that the *g* found in one sample is the same as the *g* in another sample. Research on Western samples have shown that the *g* factors extracted from different intelligence test batteries are highly correlated ($r \geq .95$), indicating that in these groups, the *g* factor is the functionally equivalent from test to test (Floyd, Reynolds, Farmer, & Kranzler, 2013; Johnson, Bouchard, Krueger, McGue, & Gottesman, 2004; Johnson, te Nijenhuis, & Bouchard, 2008; Keith, Kranzler, & Flanagan, 2001; Stauffer, Ree, & Carretta, 1996). It is unknown whether functionally equivalent *g* factors would appear in non-Western samples, though given that our results align so well with the findings from Western samples it seems plausible that many of these *g* factors could be similar.

Conclusion

Whether "intelligence" exists across cultures or whether the term has the same meaning across cultures is unknowable and probably irrelevant. The term is culturally loaded and will often have a somewhat altered meaning when translated into other languages. However, the statistical abstraction of Spearman's *g* is apparent across cultural groups in 31 non-Western, nonindustrialized nations. These results had a remarkable degree of uniformity when data sets were subjected to EFA when using the modern methods of factor selection and rotation. The resulting *g* factor explained an average of 45.9% observed variable variance, which is about the same percentage found in Western samples.

With approximately half of variance in cognitive task performance having non-*g* sources of variance, we believe that other traits may be important in explaining cognitive performance of both non-Western and Western groups. There is likely a great deal of cross-cultural variability in the level of importance of non-*g* traits, possibly because of environmental differences or because of the varying emphasis that different cultures put on developing some cognitive traits (Berry, 1966, 1986). We believe that culturally sensitive assessment and education that take into account these varying non-*g* abilities would be beneficial to individuals from many different cultures (Serpell, 2011). This does not change the fact, though, that general cognitive ability is a dominant trait in the cognition of the non-Western groups we investigated and that it is likely a universal human trait.

References

*References marked with an asterisk indicate cognitive data sources included in the analysis.

*Abdelhamid, G. S. M., Gómez-Benito, J., Abdeltawwab, A. T. M., Bakr, M. H. S. A., & Kazem, A. M. (2017). Hierarchical structure of the Wechsler Adult Intelligence Scale–fourth edition with an Egyptian sample. *Journal of Psychoeducational Assessment*. Advance online publication. http://dx.doi.org/10.1177/0734282917732857

*Agrawal, R., & Kumar, A. (1993). The relationship between intelligence and reaction time as a function of task and person variables. *Personality and Individual Differences, 14,* 287–288. http://dx.doi.org/10.1016/0191-8869(93)90207-J

Alaraudanjoki, E. (2003). *Nepalese child labourers' life-contexts, cognitive skills and well-being*. Finland: University of Jyväskylä.

Anderson, B. (1993). Evidence from the rat for a general factor that underlies cognitive performance and that relates to brain size: Intelligence? *Neuroscience Letters, 153,* 98–102. http://dx.doi.org/10.1016/0304-3940(93)90086-Z

Arden, R., & Adams, M. J. (2016). A general intelligence factor in dogs. *Intelligence, 55,* 79–85. http://dx.doi.org/10.1016/j.intell.2016.01.008

Armstrong, E. L., te Nijenhuis, J., Woodley of Menie, M. A., Fernandes, H. B. F., Must, O., & Must, A. (2016). A NIT-picking analysis: Abstractness dependence of subtests correlated to their Flynn effect magnitudes. *Intelligence, 57,* 1–6. http://dx.doi.org/10.1016/j.intell.2016.02.009

Armstrong, E. L., & Woodley, M. A. (2014). The rule-dependence model explains the commonalities between the Flynn effect and IQ gains via retesting. *Learning and Individual Differences, 29,* 41–49. http://dx.doi.org/10.1016/j.lindif.2013.10.009

*Attallah, S. E. F., Ahmed, K. Y., & Meisenberg, G. (2014). Factor structure of Wechsler Intelligence Scale for Children-third edition (WISC-III) among gifted students in Sudan. *The Mankind Quarterly, 55,* 147–170.

*Bakhiet, S. F., Albursan, I. S., Al Qudah, M. F., Abduljabbar, A. S., Aljomaa, S. S., Toto, H. S. A., & Lynn, R. (2017). Sex differences on the WISC-III among children in Sudan and the United States. *Journal of Biosocial Science, 49,* 792–797. http://dx.doi.org/10.1017/S0021932016000432

*Bangirana, P., Sikorskii, A., Giordani, B., Nakasujja, N., & Boivin, M. J. (2015). Validation of the CogState battery for rapid neurocognitive assessment in Ugandan school age children. *Child and Adolescent Psychiatry and Mental Health, 9,* 38. http://dx.doi.org/10.1186/s13034-015-0063-6

*Bau, L. P., & Dyck, M. J. (1992). Predicting the peacetime performance of military officers: Officer selection in the Papua New Guinea Defence Force. *South Pacific Journal of Psychology, 5,* 27–37. http://dx.doi.org/10.1017/S0257543400001516

*Belay, T. G. (2006). *Zinc, iron, iodine and cognitive function of women from Sidama Zone, Southern Ethiopia* (Unpublished master's thesis). Oklahoma State University, Stillwater, OK.

Bernardoni, L. C. (1964). A culture fair intelligence for the ugh, no, and oo-la-la cultures. *The Personnel and Guidance Journal, 42,* 554–557. http://dx.doi.org/10.1002/j.2164-4918.1964.tb04701.x

Berry, J. W. (1966). Temne and Eskimo perceptual skills. *International Journal of Psychology, 1,* 207–229. http://dx.doi.org/10.1080/00207596608247156

Berry, J. W. (1986). A cross-cultural view of intelligence. In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence?* (pp. 35–38). Norwood, NJ: Ablex Publishing.

Berry, J. W., & Bennett, J. A. (1992). Cree conceptions of cognitive competence. *International Journal of Psychology, 27,* 73–88. http://dx.doi.org/10.1080/00207599208246867

*Berry, J. W., van de Koppel, J. M. H., Sénéchal, C., Annis, R. C., Bahuchet, S., Cavalli-Sforza, L. L., & Witkin, H. A. (1986). *On the edge of the forest: Cultural adaptation and cognitive development in central Africa*. Berwyn, PA: Swets of North America.

Bhatia, C. M. (1955). *Performance tests of intelligence under Indian conditions*. London, UK: Oxford University Press.

*Bogale, A., Stoecker, B. J., Kennedy, T., Hubbs-Tait, L., Thomas, D., Abebe, Y., & Hambidge, K. M. (2013). Nutritional status and cognitive performance of mother-child pairs in Sidama, Southern Ethiopia. *Maternal and Child Nutrition, 9,* 274–284. http://dx.doi.org/10.1111/j.1740-8709.2011.00345.x

Booth, M. Z. (2002). Swazi concepts of intelligence: The universal versus the local. *Ethos, 30,* 376–400. http://dx.doi.org/10.1525/eth.2002.30.4.376

Boring, E. G. (1923). Intelligence as the tests test it. *New Republic, 35,* 35–36.

*Brimble, A. R. (1963). The construction of a non-verbal intelligence test in Northern Rhodesia. *Rhodes-Livingstone Journal, 34,* 23–35.

*Bulatao, J., & Reyes-Juan, I. (1968). Intellectual and creativity test factors in high school. *Philippine Journal of Psychology, 1,* 42–52.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., . . . Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour, 2,* 637–644. http://dx.doi.org/10.1038/s41562-018-0399-z

Canivez, G. L. (2016). Bifactor modeling in construct validation of multifactored tests: Implications for understanding multidimensional constructs and test interpretation. In K. Schwizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advancements* (pp. 247–271). Gottingen, Germany: Hogrefe.

Canivez, G. L., & Watkins, M. W. (2010). Investigation of the factor structure of the Wechsler Adult Intelligence Scale—fourth edition (WAIS-IV): Exploratory and higher order factor analyses. *Psychological Assessment, 22,* 827–836. http://dx.doi.org/10.1037/a0020429

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511571312

Central Intelligence Agency. (2017). *World factbook*. Washington, DC: Author. Retrieved from https://www.cia.gov/library/publications/the-world-factbook/index.html

Church, A. T., & Katigbak, M. S. (1987). Ecocultural bias in culture-specific intelligence tests in an ecologically diverse culture: The Philippines. *Intelligence, 11,* 371–389. http://dx.doi.org/10.1016/0160-2896(87)90018-3

*Church, A. T., Katigbak, M. S., & Almario-Velazco, G. (1985). Psychometric intelligence and adaptive competence in rural Philippine children. *Intelligence, 9,* 317–340. http://dx.doi.org/10.1016/0160-2896(85)90018-2

Claeys, W. (1972). The factor structure of intelligence among teachers in the Congo. In P. J. Drenth & L. J. Cronbach (Eds.), *Mental tests and cultural adaptation* (pp. 381–390). The Hague, the Netherlands: Mouton. http://dx.doi.org/10.1515/9783111556482-046

Cleary, T. A., Humphreys, L. G., Kendrick, S. A., & Wesman, A. (1975). Educational uses of tests with disadvantaged students. *American Psychologist, 30,* 15–41. http://dx.doi.org/10.1037/0003-066X.30.1.15

*Conant, L. L., Fastenau, P. S., Giordani, B. J., Boivin, M. J., Opel, B., & Dia Nseyila, D. (1999). Modality specificity of memory span tasks among Zaïrian children: A developmental perspective. *Journal of Clinical and Experimental Neuropsychology, 21,* 375–384. http://dx.doi.org/10.1076/jcen.21.3.375.912

*Condon, D. M., & Revelle, W. (2016). Selected ICAR data from the SAPA-Project: Development and initial validation of a public-domain measure. *Journal of Open Psychology Data, 4,* e1. http://dx.doi.org/10.5334/jopd.25

Conway, A. R. A., & Kovacs, K. (2018). The nature of the general factor of intelligence. In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 49–63). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/9781316817049.005

Cucina, J. M., & Howardson, G. N. (2017). Woodcock-Johnson-III, Kaufman Adolescent and Adult Intelligence Test (KAIT), Kaufman Assessment Battery for Children (KABC), and Differential Ability Scales (DAS) support Carroll but not Cattell-Horn. *Psychological Assessment, 29,* 1001–1015. http://dx.doi.org/10.1037/pas0000389

*Dasen, P. R. (1984). The cross-cultural study of intelligence: Piaget and the Baoulé. *International Journal of Psychology, 19,* 407–434. http://dx.doi.org/10.1080/00207598408247539

*Dawson, J. L. M. (1967a). Cultural and physiological influences upon spatial-perceptual processes in West Africa. Part I. *International Journal of Psychology, 2,* 115–128. http://dx.doi.org/10.1080/00207596708247208

*Dawson, J. L. M. (1967b). Cultural and physiological influences upon spatial-perceptual processes in West Africa. Part II. *International Journal of Psychology, 2,* 171–185. http://dx.doi.org/10.1080/00207596708247214

*Díaz, A., Sellami, K., Infanzón, E., Lanzón, T., & Lynn, R. (2012). A comparative study of general intelligence in Spanish and Moroccan samples. *The Spanish Journal of Psychology, 15,* 526–532. http://dx.doi.org/10.5209/rev_SJOP.2012.v15.n2.38863

Dutton, E., Becker, D., Osman, H. A., Bakhiet, S. F., Essa, Y. A. S., Ali, H. A. A., . . . Alfaleh, A. S. H. (2018). The Raven's test performance of South Sudanese samples: A validation of criticisms of the utility of Raven's among Sub-Saharan Africans. *Personality and Individual Differences, 128,* 122–126. http://dx.doi.org/10.1016/j.paid.2018.02.018

*Elwan, F. Z. (1996). Factor structure of the Kaufman Assessment Battery for Children with Egyptian schoolchildren. *Psychological Reports, 78,* 99–110. http://dx.doi.org/10.2466/pr0.1996.78.1.99

*Engle, P. L., Klein, R. E., Kagan, J., & Yarbrough, C. K. (1977). Cognitive performance during middle childhood in rural Guatemala. *The Journal of Genetic Psychology, 131,* 291–307. http://dx.doi.org/10.1080/00221325.1977.10533300

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4,* 272–299. http://dx.doi.org/10.1037/1082-989X.4.3.272

*Fernald, L. C. H., Weber, A., Galasso, E., & Ratsifandrihamanana, L. (2011). Socioeconomic gradients and child development in a very low income population: Evidence from Madagascar. *Developmental Science, 14,* 832–847. http://dx.doi.org/10.1111/j.1467-7687.2010.01032.x

Fernandes, H. B. F., Woodley, M. A., & te Nijenhuis, J. (2014). Differences in cognitive abilities among primates are concentrated on *G*: Phenotypic and phylogenetic comparisons with two meta-analytical databases. *Intelligence, 46,* 311–322. http://dx.doi.org/10.1016/j.intell.2014.07.007

Floyd, R. G., Reynolds, M. R., Farmer, R. L., & Kranzler, J. H. (2013). Are the general factors from sifferent child and adolescent intelligence tests the same? Results from a five-sample, six-test analysis. *School Psychology Review, 42,* 383–401.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101,* 171–191. http://dx.doi.org/10.1037/h0090408

Frazier, T. W., & Youngstrom, E. A. (2007). Historical increase in the number of factors measured by commercial tests of cognitive ability: Are we overfactoring? *Intelligence, 35,* 169–182. http://dx.doi.org/10.1016/j.intell.2006.07.002

Freeman, F. N. (1923). A referendum of psychologists. *The Century, 107,* 237–245.

Galsworthy, M. J., Paya-Cano, J. L., Monleón, S., & Plomin, R. (2002). Evidence for general cognitive ability (*g*) in heterogeneous stock mice and an analysis of potential confounds. *Genes Brain & Behavior, 1,* 88–95. http://dx.doi.org/10.1034/j.1601-183X.2002.10204.x

*Gauvain, M., & Munroe, R. L. (2009). Contributions of societal modernity to cognitive development: A comparison of four cultures. *Child Development, 80,* 1628–1642. http://dx.doi.org/10.1111/j.1467-8624.2009.01358.x

Gichuhi, J. (1999). *An examination of the Wechsler Intelligence Scale for Children-III for predicting performance on a national examination of grade eight children in public primary schools* (Unpublished doctoral dissertation). Biola University, La Mirada, CA.

Gottfredson, L. S. (1997a). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence, 24,* 13–23. http://dx.doi.org/10.1016/S0160-2896(97)90011-8

Gottfredson, L. S. (1997b). Why *g* matters: The complexity of everyday life. *Intelligence, 24,* 79–132. http://dx.doi.org/10.1016/S0160-2896(97)90014-3

*Grigorenko, E. L., Geissler, P. W., Prince, R., Okatcha, F., Nokes, C., Kenny, D. A., . . . Sternberg, R. J. (2001). The organisation of Luo conceptions of intelligence: A study of implicit theories in a Kenyan village. *International Journal of Behavioral Development, 25,* 367–378. http://dx.doi.org/10.1080/01650250042000348

*Grigorenko, E. L., Ngorosho, D., Jukes, M., & Bundy, D. (2006). Reading in able and disabled readers from around the world: Same or different? An illustration from a study of reading-related processes in a Swahili sample of siblings. *Journal of Research in Reading, 29,* 104–123. http://dx.doi.org/10.1111/j.1467-9817.2006.00295.x

*Gurven, M., Fuerstenberg, E., Trumble, B., Stieglitz, J., Beheim, B., Davis, H., & Kaplan, H. (2017). Cognitive performance across the life course of Bolivian forager-farmers with limited schooling. *Developmental Psychology, 53,* 160–176. http://dx.doi.org/10.1037/dev0000175

Gustafsson, J.-E., & Wolff, U. (2015). Measuring fluid intelligence at age four. *Intelligence, 50,* 175–185. http://dx.doi.org/10.1016/j.intell.2015.04.008

*Guthrie, G. M. (1963). Structure of abilities in a non-western culture. *Journal of Educational Psychology, 54,* 94–103. http://dx.doi.org/10.1037/h0040110

Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika, 19,* 149–161. http://dx.doi.org/10.1007/BF02289162

Haier, R. J. (2017). *The neuroscience of intelligence*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/9781316105771

*Haile, D., Gashaw, K., Nigatu, D., & Demelash, H. (2016). Cognitive function and associated factors among school age children in Goba Town, South-East Ethiopia. *Cognitive Development, 40,* 144–151. http://dx.doi.org/10.1016/j.cogdev.2016.09.002

Hakstian, A. R., Rogers, W. T., & Cattell, R. B. (1982). The behavior of number-of-factors rules with simulated data. *Multivariate Behavioral Research, 17,* 193–219. http://dx.doi.org/10.1207/s15327906mbr1702_3

Halpern, D. F., & Butler, H. A. (2018). Is critical thinking a better model of intelligence? In R. J. Sternberg (Ed.), *The nature of human intelligence* (pp. 183–196). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/9781316817049.013

*Hashmi, M. A., Tirmizi, S. H., & Shah, A. F. (2010). Development and validation of intelligence test for grade six students (age group 11–12 years). *The Journal of Educational Research, 13,* 19–29.

*Hashmi, M. A., Tirmizi, S. H., Shah, A. F., & Khan, T. M. (2011). Development and validation of intelligence test for grade seven students (age group 12–13 years). *The Journal of Educational Research, 14,* 6–14.

*Hendrawan, D. (2013). Factors affecting letter fluency performance among Indonesian undergraduate students: A short report. *Procedia: Social and Behavioral Sciences, 97,* 266–270. http://dx.doi.org/10.1016/j.sbspro.2013.10.232

Herndon, J. G., Moss, M. B., Rosene, D. L., & Killiany, R. J. (1997). Patterns of cognitive decline in aged rhesus monkeys. *Behavioural Brain Research, 87,* 25–34. http://dx.doi.org/10.1016/S0166-4328(96)02256-5

Herrnstein, R. J., & Murray, C. (1996). *The bell curve: Intelligence and class structure in American life* (2nd ed.). New York, NY: Free Press.

Hill, W. D., Davies, G., Liewald, D. C., Payton, A., McNeil, C. J., Whalley, L. J., . . . Deary, I. J. (2016). Examining non-syndromic autosomal recessive intellectual disability (NS-ARID) genes for an enriched association with intelligence differences. *Intelligence, 54,* 80–89. http://dx.doi.org/10.1016/j.intell.2015.11.005

Hodos, W., & Campbell, C. B. G. (1969). *Scala naturae*: Why there is no theory in comparative psychology. *Psychological Review, 76,* 337–350. http://dx.doi.org/10.1037/h0027523

*Holding, P., Anum, A., van de Vijver, F. J. R., Vokhiwa, M., Bugase, N., Hossen, T., . . . Gomes, M. (2018). Can we measure cognitive constructs consistently within and across cultures? Evidence from a test battery in Bangladesh, Ghana, and Tanzania. *Applied Neuropsychology Child, 7,* 1–13. http://dx.doi.org/10.1080/21622965.2016.1206823

Hopkins, W. D., Russell, J. L., & Schaeffer, J. (2014). Chimpanzee intelligence is heritable. *Current Biology, 24,* 1649–1652. http://dx.doi.org/10.1016/j.cub.2014.05.076

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179–185. http://dx.doi.org/10.1007/BF02289447

*Humble, S., & Dixon, P. (2017). The effects of schooling, family and poverty on children's attainment, potential and confidence—Evidence from Kinondoni, Dar es Salaam, Tanzania. *International Journal of Educational Research, 83,* 94–106. http://dx.doi.org/10.1016/j.ijer.2017.03.001

Hunt, E. (2011). *Human intelligence*. New York, NY: Cambridge University Press.

*Irvine, S. H. (1964). *A psychological study of selection problems at the end of primary schooling in Southern Rhodesia* (Unpublished doctoral dissertation). University of London, London, UK.

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

Johnson, W., Bouchard, T. J., Jr., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one *g*: Consistent results from three test batteries. *Intelligence, 32,* 95–107. http://dx.doi.org/10.1016/S0160-2896(03)00062-X

Johnson, W., te Nijenhuis, J., & Bouchard, T. J., Jr. (2008). Still just 1 *g*: Consistent results from five test batteries. *Intelligence, 36,* 81–95. http://dx.doi.org/10.1016/j.intell.2007.06.001

Jones, B. D., Rakes, L., & Landon, K. (2013). Malawian secondary students' beliefs about intelligence. *International Journal of Psychology, 48,* 785–796. http://dx.doi.org/10.1080/00207594.2012.716906

Kagan, J., Klein, R. E., Finley, G. E., Rogoff, B., Nolan, E., & Greenbaum, C. (1979). *A cross-cultural study of cognitive development*. Illinois: University of Chicago Press. http://dx.doi.org/10.2307/1165916

Kathuria, R., & Serpell, R. (1998). Standardization of the Panga Munthu Test—A nonverbal cognitive test developed in Zambia. *The Journal of Negro Education, 67,* 228–241. http://dx.doi.org/10.2307/2668192

Kearney, G. E. (1966). *Orokaiva papers: Miscellaneous papers on the Orokaiva of north east Papua* (New Guinea Research Bulletin No. 13). Canberra, Australia: New Guinea Research Unit, The Australian National University.

Keith, T. Z., Kranzler, J. H., & Flanagan, D. P. (2001). What does the Cognitive Assessment System (CAS) measure? Joint confirmatory factor analysis of the CAS and the Woodcock-Johnson Tests of Cognitive Ability (3rd ed.). *School Psychology Review, 30,* 89–119.

Kwate, N. O. A. (2001). Intelligence or misorientation? Eurocentrism in the WISC-III. *The Journal of Black Psychology, 27,* 221–238. http://dx.doi.org/10.1177/0095798401027002005

Larsen, R., & Warne, R. T. (2010). Estimating confidence intervals for eigenvalues in exploratory factor analysis. *Behavior Research Methods, 42,* 871–876. http://dx.doi.org/10.3758/BRM.42.3.871

*Lean, G., & Clements, M. A. (1981). Spatial ability, visual imagery, and mathematical performance. *Educational Studies in Mathematics, 12,* 267–299. http://dx.doi.org/10.1007/BF00311060

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904). "General intelligence," objectively determined and measured. *Journal of Personality and Social Psychology, 86,* 96–111. http://dx.doi.org/10.1037/0022-3514.86.1.96

Lubinski, D., & Humphreys, L. G. (1992). Some bodily and medical correlates of mathematical giftedness and commensurate levels of socioeconomic status. *Intelligence, 16,* 99–115. http://dx.doi.org/10.1016/0160-2896(92)90027-O

Lynn, R. (2010). The average IQ of sub-Saharan Africans assessed by the Progressive Matrices: A reply to Wicherts, Dolan, Carlson & van der Maas. *Learning and Individual Differences, 20,* 152–154. http://dx.doi.org/10.1016/j.lindif.2010.03.009

Lynn, R. (2015). *Race differences in intelligence: An evolutionary analysis* (2nd ed.). Arlington, VA: Washington Summit Publishers.

Lynn, R., & Owen, K. (1994). Spearman's hypothesis and test score differences between whites, Indians, and blacks in South Africa. *The Journal of General Psychology, 121,* 27–36. http://dx.doi.org/10.1080/00221309.1994.9711170

MacArthur, R. S., Irvine, S. H., & Brimble, A. R. (1964). *The Northern Rhodesia mental ability survey 1963*. Lusaka, Zambia: Rhodes-Livingstone Institute.

*Mahmood, S. (2013). *Development and validation of an indigenous intelligence test* (Unpublished doctoral dissertation). University of Peshawar, Peshawar, Pakistan.

Matzel, L. D., & Sauce, B. (2017). Individual differences: Case studies of rodent and primate intelligence. *Journal of Experimental Psychology: Animal Learning and Cognition, 43,* 325–340. http://dx.doi.org/10.1037/xan0000152

*McCoy, D. C., Zuilkowski, S. S., Yoshikawa, H., & Fink, G. (2017). Early childhood care and education and school readiness in Zambia. *Journal of Research on Educational Effectiveness, 10,* 482–506. http://dx.doi.org/10.1080/19345747.2016.1250850

*McFie, J. (1954). African performance on an intelligence test. *Uganda Journal, 18,* 34–43.

*McFie, J. (1961). The effect of education on African performance on a group of intellectual tests. *The British Journal of Educational Psychology, 31,* 232–240. http://dx.doi.org/10.1111/j.2044-8279.1961.tb01712.x

Meehl, P. E. (2006). The power of quantitative thinking. In N. G. Waller, L. J. Yonce, W. M. Grove, D. Faust, & M. F. Lenzenweger (Eds.), *A Paul Meehl reader: Essays on the practice of scientific psychology* (pp. 433–444). Mahwah, NJ: Erlbaum.

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science, 11,* 512–534. http://dx.doi.org/10.1177/1745691616635612

*Miezah, D. (2015). *Validation of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV) in the Ghanian population* (Unpublished master's thesis). University of Ghana, Accra, Ghana.

Miron, M. (1975). A study of cross-cultural factorial structure of intelligence. *Psychologia, 18,* 92–94.

*Mourgues, C., Tan, M., Hein, S., Ojanen, E., Reich, J., Lyytinen, H., & Grigorenko, E. L. (2016). Paired associate learning tasks and their contribution to reading skills. *Learning and Individual Differences, 46,* 54–63. http://dx.doi.org/10.1016/j.lindif.2014.12.003

Mpofu, E. (2004). Being intelligent with Zimbabweans. In R. J. Sternberg (Ed.), *International handbook of intelligence* (pp. 364–390). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511616648.014

Mpofu, E., Ntinda, K., & Oakland, T. (2012). Understanding human abilities in Sub-Saharan African settings. *Online Readings in Psychology and Culture*. Advance online publication. http://dx.doi.org/10.9707/2307-0919.1036

*Mwaura, P. A. M., Sylva, K., & Malmberg, L.-E. (2008). Evaluating the Madrasa preschool programme in East Africa: A quasi-experimental study. *International Journal of Early Years Education, 16,* 237–255. http://dx.doi.org/10.1080/09669760802357121

Ngara, C., & Porath, M. (2004). Shona culture of Zimbabwe's views of giftedness. *High Ability Studies, 15,* 189–209. http://dx.doi.org/10.1080/1359813042000314772

Nsamenang, A. B. (2006). Human ontogenesis: An indigenous African view on development and intelligence. *International Journal of Psychology, 41,* 293–297. http://dx.doi.org/10.1080/00207590544000077

Oakland, T., Douglas, S., & Kane, H. (2016). Top ten standardized tests used internationally with children and youth by school psychologists in 64 countries: A 24-year follow-up study. *Journal of Psychoeducational Assessment, 34,* 166–176. http://dx.doi.org/10.1177/0734282915595303

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test.

Behavior Research Methods, Instruments, & Computers, 32, 396–402. http://dx.doi.org/10.3758/BF03200807

*O'Donnell, K., Murphy, R., Ostermann, J., Masnick, M., Whetten, R. A., Madden, E., . . . Whetten, K. (2012). A brief assessment of learning for orphaned and abandoned children in low and middle income countries. *AIDS and Behavior, 16,* 480–490. http://dx.doi.org/10.1007/s10461-011-9940-z

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349,* aac4716-4711.

*Orbell, S. F. (1981). Formal operational thinking among African adolescents. *Journal of Instructional Psychology, 8,* 10–14.

Ord, I. G. (1970). *Mental test for pre-literates: Resulting mainly from New Guinea studies*. Melbourne, Australia: The Jacaranda Press.

*Panza Lombardo, M. A. (2016). Transferencia y permanencia del entrenamiento *dual n-back* en la inteligencia fluida y en la memoria de trabajo verbal [Transfer and permanence of dual n-back training on fluid intelligence and verbal working memory]. *Psicodebate, 16,* 49–82. http://dx.doi.org/10.18682/pd.v16i1.566

Penke, L., Maniega, S. M., Bastin, M. E., Valdés Hernández, M. C., Murray, C., Royle, N. A., . . . Deary, I. J. (2012). Brain white matter tract integrity as a neural foundation for general intelligence. *Molecular Psychiatry, 17,* 1026–1030. http://dx.doi.org/10.1038/mp.2012.66

*Peters, E., Baker, D. P., Dieckmann, N. F., Leon, J., & Collins, J. (2010). Explaining the effect of education on health: A field study in Ghana. *Psychological Science, 21,* 1369–1376. http://dx.doi.org/10.1177/0956797610381506

Pietschnig, J., Penke, L., Wicherts, J. M., Zeiler, M., & Voracek, M. (2015). Meta-analysis of associations between human brain volume and intelligence differences: How strong are they and what do they mean? *Neuroscience and Biobehavioral Reviews, 57,* 411–432. http://dx.doi.org/10.1016/j.neubiorev.2015.09.017

*Pitchford, N. J., & Outhwaite, L. A. (2016). Can touch screen tablets be used to assess cognitive and motor skills in early years primary school children? A cross-cultural study. *Frontiers in Psychology, 7,* 1666. http://dx.doi.org/10.3389/fpsyg.2016.01666

Plomin, R. (2001). The genetics of *g* in human and mouse. *Nature Reviews Neuroscience, 2,* 136–141. http://dx.doi.org/10.1038/35053584

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2016). Top 10 replicated findings from behavioral genetics. *Perspectives on Psychological Science, 11,* 3–23. http://dx.doi.org/10.1177/1745691615617439

*Pongcharoen, T., DiGirolamo, A. M., Ramakrishnan, U., Winichagoon, P., Flores, R., & Martorell, R. (2011). Long-term effects of iron and zinc supplementation during infancy on cognitive function at 9 y of age in northeast Thai children: A follow-up study. *The American Journal of Clinical Nutrition, 93,* 636–643. http://dx.doi.org/10.3945/ajcn.110.002220

Porteus, S. D. (1965). *Porteus maze test: Fifty years' application*. Palo Alto, CA: Pacific Books.

Proctor, B. E., Kranzler, J. H., Rosenbloom, A. L., Martinez, V., & Guevara-Aguire, J. (2000). An initial investigation of validation of the matrix analogies test-expanded form in Ecuador. *Psychological Reports, 86,* 445–453. http://dx.doi.org/10.2466/pr0.2000.86.2.445

Protzko, J. (2017). Effects of cognitive training on the structure of intelligence. *Psychonomic Bulletin & Review, 24,* 1022–1031. http://dx.doi.org/10.3758/s13423-016-1196-1

*Rasheed, M. A., Pham, S., Memon, U., Siyal, S., Obradović, J., & Yousafzai, A. K. (2017). Adaptation of the Wechsler Preschool and Primary Scale of Intelligence-III and lessons learned for evaluating intelligence in low-income settings. *International Journal of School & Educational Psychology, 6,* 197–207. http://dx.doi.org/10.1080/21683603.2017.1302851

*Rehna, T., & Hanif, R. (2017). Verbal cognitive abilities and emotional and behavioral problems of secondary school children. *Foundational University Journal of Psychology, 2,* 20–35.

Reyes, A., Pacifico, R., Benitez, B., Villanueva-Uy, E., Lam, H., & Ostrea, E. M., Jr. (2010). Use of the Griffiths Mental Development Scales in an agro-industrial province in the Philippines. *Child: Care, Health and Development, 36,* 354–360. http://dx.doi.org/10.1111/j.1365-2214.2010.01080.x

Richardson, K. (2002). What IQ tests test. *Theory & Psychology, 12,* 283–314. http://dx.doi.org/10.1177/0959354302012003012

Rindermann, H. (2013). African cognitive ability: Research, results, divergences and recommendations. *Personality and Individual Differences, 55,* 229–233. http://dx.doi.org/10.1016/j.paid.2012.06.022

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science, 29,* 1358–1369. http://dx.doi.org/10.1177/0956797618774253

Rothhammer, F., & Llop, E. (1976). Amerindian descent and intellectual performance in Chilean university students. *Human Biology, 48,* 455–464.

*Ruffieux, N., Njamnshi, A. K., Mayer, E., Sztajzel, R., Eta, S. C., Doh, R. F., . . . Hauert, C. A. (2010). Neuropsychology in Cameroon: First normative data for cognitive tests among school-aged children. *Child Neuropsychology, 16,* 1–19. http://dx.doi.org/10.1080/09297040902802932

Rushton, J. P., & Ankney, C. D. (2009). Whole brain size and general mental ability: A review. *The International Journal of Neuroscience, 119,* 691–731. http://dx.doi.org/10.1080/00207450802325843

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika, 22,* 53–61. http://dx.doi.org/10.1007/BF02289209

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist, 50,* 138–166. http://dx.doi.org/10.1080/00461520.2015.1036274

*Sen, A., Jensen, A. R., Sen, A. K., & Arora, I. (1983). Correlation between reaction time and intelligence in psychometrically similar groups in America and India. *Applied Research in Mental Retardation, 4,* 139–152. http://dx.doi.org/10.1016/0270-3092(83)90006-1

Serpell, R. (2011). Social responsibility as a dimension of intelligence, and as an educational goal: Insights from programmatic research in an African society. *Child Development Perspectives, 5,* 126–133. http://dx.doi.org/10.1111/j.1750-8606.2011.00167.x

Serpell, R., & Jere-Folotiya, J. (2008). Developmental assessment, cultural context, gender, and schooling in Zambia. *International Journal of Psychology, 43,* 88–96. http://dx.doi.org/10.1080/00207590701859184

*Shrestha, R. M. (1994). *Effect of iodine and iron supplementation on physical, psychomotor and mental development in primary school children in Malawi* (Unpublished doctoral dissertation). Wageningen University, Wageningen, the Netherlands.

Singham, M. (1995). Race and intelligence: What are the issues? *Phi Delta Kappan, 77,* 271–278.

*Songy, D. G. (2007). Predicting success in academic achievement of major seminarians in Papua New Guinea: A comparison of cognitive test results and grade point averages. *Contemporary PNG Studies, 7,* 59–71.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology, 15,* 201–293. http://dx.doi.org/10.2307/1412107

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

Spitz, H. H. (1978). The universal nature of human intelligence: Evidence from games. *Intelligence, 2,* 371–379. http://dx.doi.org/10.1016/0160-2896(78)90026-0

Srivastava, A. K., & Misra, G. (2001). Lay people's understanding and use of intelligence: An Indian perspective. *Psychology and Developing Societies, 13,* 25–49. http://dx.doi.org/10.1177/097133360101300102

Stauffer, J. M., Ree, M. J., & Carretta, T. R. (1996). Cognitive-components tests are not much more than *g*: An extension of Kyllonen's analyses. *The Journal of General Psychology, 123,* 193–205. http://dx.doi.org/10.1080/00221309.1996.9921272

Stemler, S. E., Chamvu, F., Chart, H., Jarvin, L., Jere, J., Hart, L., & Grigorenko, E. L. (2009). Assessing competencies in reading and mathematics in Zambian children. In E. L. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 157–185). New York, NY: Springer.

*Sternäng, O., Lövdén, M., Kabir, Z. N., Hamadani, J. D., & Wahlin, Å. (2016). Different context but similar cognitive structures: Older adults in rural Bangladesh. *Journal of Cross-Cultural Gerontology, 31,* 143–156. http://dx.doi.org/10.1007/s10823-016-9284-2

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511509612

Sternberg, R. J., & Detterman, D. K. (Eds.). (1986). *What is intelligence?* Norwood, NJ: Ablex Publishing.

Sternberg, R. J., & Grigorenko, E. L. (2001). Ability testing across cultures. In L. A. Suzuki, J. G. Ponterotto, & P. J. Meller (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (2nd ed., pp. 335–359). San Francisco, CA: Jossey-Bass.

*Sternberg, R. J., Grigorenko, E. L., Ngorosho, D., Tantufuye, E., Mbise, A., Nokes, C., . . . Bundy, D. A. (2002). Assessing intellectual potential in rural Tanzanian school children. *Intelligence, 30,* 141–162. http://dx.doi.org/10.1016/S0160-2896(01)00091-5

*Sternberg, R. J., Nokes, C., Geissler, P. W., Prince, R., Okatcha, F., Bundy, D. A., & Grigorenko, E. L. (2001). The relationship between academic and practical intelligence: A case study in Kenya. *Intelligence, 29,* 401–418. http://dx.doi.org/10.1016/S0160-2896(01)00065-4

*Sukhatunga, K., Phattharayuttawat, S., Kowasint, C., Chantra, J., Bunnagulrote, W., Bunnagulrote, K., & Imaroonrak, S. (2002). Preliminary validation of the Thai version of the Wechsler Abbreviated Scale of Intelligence. *Siriraj Medical Journal, 54,* 345–353.

*Tan, M., Reich, J., Hart, L., Thuma, P. E., & Grigorenko, E. L. (2014). Examining the specific effects of context on adaptive behavior and achievement in a rural African community: Six case studies from rural areas of Southern province, Zambia. *Journal of Autism and Developmental Disorders, 44,* 271–282. http://dx.doi.org/10.1007/s10803-012-1487-y

Terman, L. M. (1921). Intelligence and its measurement: A symposium—II. *Journal of Educational Psychology, 12,* 127–133. http://dx.doi.org/10.1037/h0064940

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10694-000

Tsukahara, J. S., Harrison, T. L., & Engle, R. W. (2016). The relationship between baseline pupil size and intelligence. *Cognitive Psychology, 91,* 109–123. http://dx.doi.org/10.1016/j.cogpsych.2016.10.001

United Nations. (2018). *World economic situation and prospects*. New York, NY: Author. Retrieved from https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/publication/WESP2018_Full_Web-1.pdf

*van den Briel, T., West, C. E., Bleichrodt, N., van de Vijver, F. J. R., Ategbo, E. A., & Hautvast, J. G. A. J. (2000). Improved iodine status is associated with improved mental performance of schoolchildren in Benin. *The American Journal of Clinical Nutrition, 72,* 1179–1185. http://dx.doi.org/10.1093/ajcn/72.5.1179

van der Eijk, C., & Rose, J. (2015). Risky business: Factor analysis of survey data - assessing the probability of incorrect dimensionalisation. *PLoS ONE, 10,* e0118900. http://dx.doi.org/10.1371/journal.pone.0118900

*van de Vijver, F. J. R. (2002). Inductive reasoning in Zambia, Turkey, and the Netherlands: Establishing cross-cultural equivalence. *Intelligence, 30,* 313–351. http://dx.doi.org/10.1016/S0160-2896(02)00084-3

*Veena, S. R. (2014). *Cognitive performance during childhood and early adolescence in India: Relationships to birth size, maternal nutrition during pregnancy and postnatal growth* (Unpublished doctoral dissertation). University of Southampton, Southampton, UK.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41,* 321–327. http://dx.doi.org/10.1007/BF02293557

Velicer, W. F., Eaton, C. A., & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Eds.), *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy* (pp. 41–71). New York, NY: Kluwer Academic/Plenum Press Publishers. http://dx.doi.org/10.1007/978-1-4615-4397-8_3

Vernon, P. E. (1961). *The structure of human abilities* (2nd ed.). London, UK: Methuen.

Vernon, P. E. (1965a). Environmental handicaps and intellectual development: Part I. *The British Journal of Educational Psychology, 35,* 9–20. http://dx.doi.org/10.1111/j.2044-8279.1965.tb01782.x

Vernon, P. E. (1965b). Environmental handicaps and intellectual development: Part II. *The British Journal of Educational Psychology, 35,* 117–126. http://dx.doi.org/10.1111/j.2044-8279.1965.tb01796.x

Vernon, P. E. (1969). *Intelligence and cultural environment*. London, UK: Methuen.

*Warburton, F. W. (1951). The ability of the Gurkha recruit. *The British Journal of Psychology General Section, 42,* 123–133. http://dx.doi.org/10.1111/j.2044-8295.1951.tb00284.x

Warne, R. T. (2016). Five reasons to put the *g* back into giftedness: An argument for applying the Cattell–Horn–Carroll theory of intelligence to gifted education research and practice. *Gifted Child Quarterly, 60,* 3–15. http://dx.doi.org/10.1177/0016986215605360

Warne, R. T., & Larsen, R. (2014). Evaluating a proposed modification of the Guttman rule for determining the number of factors in an exploratory factor analysis. *Psychological Test and Assessment Modeling, 56,* 104–123.

Wicherts, J. M., Dolan, C. V., Carlson, J. S., & van der Maas, H. L. J. (2010a). Raven's test performance of sub-Saharan Africans: Average performance, psychometric properties, and the Flynn Effect. *Learning and Individual Differences, 20,* 135–151. http://dx.doi.org/10.1016/j.lindif.2009.12.001

Wicherts, J. M., Dolan, C. V., Carlson, J. S., & van der Maas, H. L. J. (2010b). Another failure to replicate Lynn's estimate of the average IQ of sub-Saharan Africans. *Learning and Individual Differences, 20,* 155–157. http://dx.doi.org/10.1016/j.lindif.2010.03.010

Wicherts, J. M., Dolan, C. V., & van der Maas, H. L. J. (2010). A systematic literature review of the average IQ of sub-Saharan Africans. *Intelligence, 38,* 1–20. http://dx.doi.org/10.1016/j.intell.2009.05.002

Wolff, H.-G., & Preising, K. (2005). Exploring item and higher order factor structure with the Schmid-Leiman solution: Syntax codes for SPSS and SAS. *Behavior Research Methods, 37,* 48–58. http://dx.doi.org/10.3758/BF03206397

Woodley, M. A., te Nijenhuis, J., Must, O., & Must, A. (2014). Controlling for increased guessing enhances the independence of the Flynn effect from *g*: The return of the Brand effect. *Intelligence, 43,* 27–34. http://dx.doi.org/10.1016/j.intell.2013.12.004

Woodley of Menie, M. A., Piffer, D., Peñaherrera, M. A., & Rindermann, H. (2016). Evidence of contemporary polygenic selection on the Big G of national cognitive ability: A cross-cultural sociogenetic analysis. *Personality and Individual Differences, 102,* 90–97. http://dx.doi.org/10.1016/j.paid.2016.06.054

Yang, S.-Y., & Sternberg, R. J. (1997). Taiwanese Chinese people's conceptions of intelligence. *Intelligence, 25,* 21–36. http://dx.doi.org/10.1016/S0160-2896(97)90005-2

Yerkes, R. M. (1921). *Psychological examining in the United States army*. Washington, DC: U.S. Government Printing Office.

Zwick, W. R., & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research, 17,* 253–269. http://dx.doi.org/10.1207/s15327906mbr1702_5

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432–442. http://dx.doi.org/10.1037/0033-2909.99.3.432

Received April 23, 2018
Revision received October 18, 2018
Accepted November 19, 2018 ■