Acta Psychologica 240 (2023) 104015

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

International differences in the speed of cognitive development: A systematic examination of the existence of the Simber Effect

David Becker [a,*], Gerhard Meisenberg [b], Edward Dutton [c], Salaheldin Farah Attallah Bakhiet [d], Fayez Abdulaziz Alfayez [e], Yossry A.S. Essa [f]

[a] *Department of Psychology, Chemnitz University of Technology, Germany*
[b] *Ulster Institute for Social Research, London, United Kingdom*
[c] *ASBiRO University, Łódź, Poland*
[d] *King Saud University, Department of Special Education, College of Education, Riyadh, Saudi Arabia*
[e] *King Saud University, Educational Administration Department, College of Education, Riyadh, Saudi Arabia*
[f] *Arish University, Department of Special Education, College of Education, Egypt*

ARTICLE INFO

Keywords:
Simber Effect
Cognitive development
Raven's matrices
meta-analysis
Cross-national

ABSTRACT

The Simber Effect refers to the phenomenon whereby, in Arabic countries, young children have an IQ that is little different from that of Western children but that these differences increase throughout childhood culminating in a difference of around 20 points by adulthood. The true nature of this phenomenon is revealed by an examination of 125 samples from all around the globe measured with Raven's Progressive Matrices. We show that in many cases different speeds of cognitive development increase the IQ score differences between countries mostly between 4 and 9 years of age, and that these increases can in part be explained by poor environmental conditions. However, the patterns are not completely clear, either in terms of regularity or strengths. Methodological problems, in particular the cross-sectional designs of the included samples, as well as the significance of the Simber Effect for country comparisons in intelligence are discussed.

1. Introduction

Bakhiet et al. (2018) recently presented a substantial body of evidence for a phenomenon they have termed the "Simber Effect."[1] The simber, known as the black stork in English, is a migratory bird found in Sudan, among other countries. Its arrival heralds a period of increased fertility in the land. However, as noted by Bakhiet et al., this fertile period is relatively brief, and the land soon reverts to its usual arid conditions. In alignment with this metaphor, Bakhiet et al. conducted a systematic literature review encompassing all known studies involving Arab children's performance on the Raven's Progressive Matrices. Their findings suggest that Arab children start their cognitive journey in a relatively 'intellectually fertile' state. At around the age of 5, IQ scores in these countries are only slightly lower than those in European countries such as Sweden, Britain, and Croatia. Nevertheless, throughout childhood, the average IQ of Arab children experiences a decline in comparison to their European counterparts. By early adulthood, the IQ of Arab children is approximately 20 points lower than that of Europeans. The authors also observe that this decline follows a relatively linear pattern year by year. As Arab children grow older, their IQ scores progressively decrease relative to the European benchmark.

Bakhiet et al. (2018) put forth several potential explanations for this phenomenon. They assert that the education systems in Arab and other developing countries are more traditionalist and didactic, with less emphasis on stimulating analytical thinking—a crucial skill required to perform well on intelligence tests, including Raven's Progressive Matrices. This holds particular significance as Raven's tests are frequently used in cross-cultural studies (Flynn, 2012). Consequently, the capacity to engage in this analytical thinking, which plays a pivotal role in the Flynn Effect (the steady increase in average IQ scores across the twentieth century in Western countries), accumulates progressively in European countries. This results in European children improving their skills in this aspect year by year, leading to a widening IQ score gap between them and Arab children. The authors note a similarity between

---

* Corresponding author at: Department of Psychology, Chemnitz University of Technology, Wilhelm-Raabe-Str. 43, D-09107 Chemnitz, Germany.
E-mail addresses: david.becker@psychologie.tu-chemnitz.de (D. Becker), e.c.dutton@dunelm.org (E. Dutton), bakhiet@ksu.edu.sa (S.F.A. Bakhiet), falfayez@ksu.edu.sa (F.A. Alfayez).

[1] In the name of this bird, a research group founded in Sudan by the educational psychologist Omar Haroon Al-Khaleefa, this group concerned with intelligence, giftedness, and indigenization of psychology.

https://doi.org/10.1016/j.actpsy.2023.104015
Received 25 January 2023; Received in revised form 13 August 2023; Accepted 25 August 2023

this situation and how IQ score disparities related to social class in the UK increase throughout childhood. Higher socioeconomic status children become immersed in an ever-growing cycle of activities that promote analytical thinking. However, even if this explanation for the Simber Effect proves valid, it prompts the need for more comprehensive research to identify the precise factors contributing to this effect. A second explanation proposed by Bakhiet et al. is that Arabs might follow a faster "Life History Strategy" (as discussed in Figueredo et al., 2013) compared to Europeans. This divergence might arise from a combination of environmental and possibly genetic factors. This suggests that Arabs mature more rapidly and initially show intellectual proximity to Europeans at a young age, only to gradually diverge as they grow older. Nonetheless, this hypothesis also raises questions about the specific genetic and environmental components—assuming that the Life History Strategy is influenced by the environment—that play a role.

The hypothesis that impoverished environmental conditions can result in a "cumulative deficit" throughout cognitive development is not a novel concept. Arthur Jensen (1974a) originally introduced this idea to account for lower cognitive test scores among African Americans. He later presented evidence showcasing the decelerated cognitive development of black children in rural Georgia, as opposed to the relatively minor effect observed in California, where conditions for black children were more favorable than in rural Georgia (Jensen, 1977). Jensen's conclusion was that the observed cumulative deficit could be primarily or entirely attributed to the detrimental impact of poor environmental conditions.

Bakhiet et al. (2018)'s "Simber Effect" bears a striking resemblance to Jensen's hypothesis. Nevertheless, the scope of the papers examined by Bakhiet et al. (2018) imposes limitations on the conclusions that can be drawn from their findings. All of these studies adopted a cross-sectional research design, and it remains unclear to what extent the disparities between age groups truly reflect genuine cognitive growth rates, as would be discerned in longitudinal research approaches. Another constraint stems from the fact that these investigations are exclusively centered on the Arab world. Consequently, it remains unresolved whether the Simber Effect is a unique cultural-spatial phenomenon or a universally applicable principle. This inquiry holds significant relevance in the context of studying disparities in intelligence across countries. The data aggregated in Lynn and Vanhanen (2002, 2006, 2012)'s works as well as Lynn and Becker (2019)'s study unveil substantial global variations in intelligence levels among nations. Using Britain as a benchmark, Lynn and Becker (2019) themselves identified negative correlations between the age of non-British samples and their IQ scores, ranging from $r = -.07$ to $-.21$ at the sample level and $r = -.17$ at the national level. Although these correlations are relatively modest and only partially significant, they suggest that beyond Britain, older adolescents and adults tend to exhibit lower IQ scores than younger children when normed against British standards. A systematic Simber Effect, if proven to be a global principle, potentially triggered by adverse living conditions, educational disparities, or even genetic factors, could thus result in an miscalculation of adult IQ scores in developing countries. This is particularly pertinent if their IQ estimations are extrapolated from samples of young children.

The primary objective of this study is to expand upon the research conducted by Bakhiet et al. (2018) by investigating the global extent of the Simber Effect and its potential underlying causes, such as impoverished living conditions or inadequate educational systems. To achieve this aim, we undertook a meta-analysis examining the pace of cognitive development from childhood to early adulthood across samples from countries spanning various levels of economic development. Our hypothesis centers on the idea that discrepancies in IQ scores among countries are partially influenced by either delayed or accelerated cognitive development. Should this be the case, these disparities in IQ scores should vary depending on the specific age group under consideration. For instance, if two countries exhibit distinct average scores while simultaneously showcasing dissimilar rates of cognitive development, the mean scores of these two nations would either diverge or converge as individuals age. Beyond ascertaining the presence or absence of the Simber Effect in worldwide samples, our secondary objective revolves around investigating whether the Simber Effect distorts the assessment of IQ score disparities between countries.

2. Method

2.1. Re-defining the Simber Effect

Bakhiet et al. (2018) introduced the concept of the Simber Effect (SE) based on their research focusing on Arab-speaking countries. Firstly, it is imperative to establish a universally applicable definition that encapsulates this underlying phenomenon, enabling its examination on a global scale. In Fig. 1 of their study, Bakhiet et al. (2018) illustrated that Arab countries achieved an average IQ score of 92 on the British IQ scale at the age of 5, which decreased to approximately 74 by the age of 18. This divergence exhibited a curvilinear or U-shaped trajectory. In contrast, as depicted by the authors in the same figure for Estonia, Croatia, and Sweden, other countries displayed average scores below 100 on the British IQ scale at age 5, but above 100 at ages 14 to 16. Essentially, their findings illustrate an age-related shift in the mean IQ score of a nation in relation to the British standard of 100. In the ensuing discussion, we will refer to a negative Simber Effect (SE↓) when this divergence extends the IQ score difference downward (as observed in Arab countries). Similarly, a positive Simber Effect (SE↑) will be utilized when this divergence extends the IQ score difference upward (as observed in the other three countries). It is crucial to differentiate between these two types of effects and an anti-Simber Effect (ASE), which entails an age-related convergence of IQ scores in a negative direction (ASE↓) or in a positive direction (ASE↑). Fig. 1 provides a graphical representation of these definitions and their potential combinations.

2.2. Data for intelligence

2.2.1. Initial considerations

As we delve into a relatively novel and unexplored phenomenon, it is imperative to minimize potential disruptive factors from the outset. Consequently, we made the initial decision to exclusively compare the outcomes of a suite of intelligence tests sharing similar structures and domains of intelligence, with a strong focus on the general factor of intelligence (*g*). This approach sought to avoid capturing specific factors and, instead, concentrate on the overarching concept. This decision was well-founded, considering that studies spanning numerous non-Western and non-industrialized nations have demonstrated the cross-cultural applicability of intelligence models with a prevailing general factor (Warne & Burningham, 2019). Neuroanatomical correlations further suggest the likely universality and biological basis of this general factor (Barbey et al., 2012; Colom et al., 2006; Haier et al., 2009; Posner & Barbey, 2020). Moreover, the selected tests should possess minimal scholastic content and language bias to ensure the comparability of results across countries. These criteria are particularly crucial for developing countries, which are of primary interest to our study. In light of these requirements, we decided to focus exclusively on the various versions of the Raven's Progressive Matrices (RPM), which include the Standard Progressive Matrices (SPM), the Colored Progressive Matrices (CPM), and the Standard Progressive Matrices Plus (SPM+) (e.g., Raven, 1981, 2000, 2008a, 2008b; Raven et al., 1985, 1998). These tests encompass matrix reasoning tasks designed to evaluate perception-bound logical thinking, fluid intelligence, or fluid reasoning (Lohman & Hagen, 2002, p.94; Petermann & Petermann, 2014; Weiß & Osterland, 1980, 1997, 2013). Notably, these tests are well-suited for assessing the general intelligence factor *g* (Jensen, 1980, 1993 p.186; Spearman, 1946, p.127; Vernon & Parry, 1949, p.234) and are commonly considered relatively 'culture fair' (Anderson et al., 1968; Jensen, 1972, pp. 81, 184, 195; Jensen, 1974b; Stacey & Carleton, 1955). However, it is

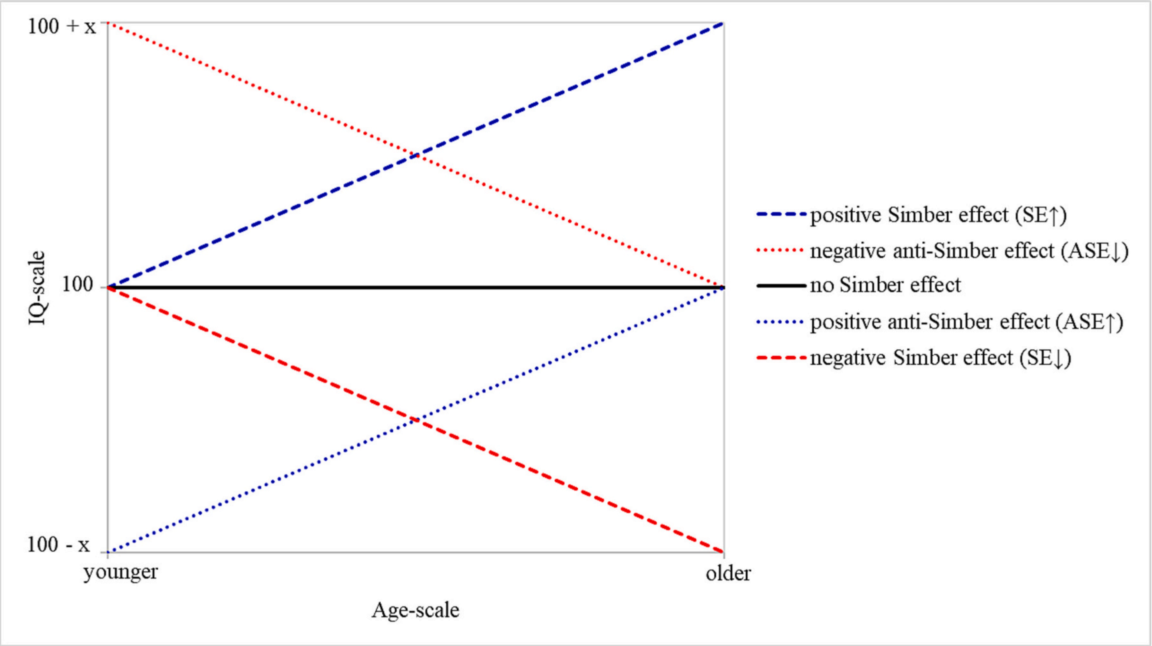

Fig. 1. Graphical abstractions of four types of Simber effects.

essential to acknowledge that recent doubts have emerged regarding the genuine cultural fairness of these tests, which prompts a nuanced view of our findings (Becker et al., 2022; Dutton et al., 2018; Fox & Mitchum, 2013). On the other hand, these tests have consistently demonstrated high reliability when administered to non-Western samples (e.g., Bakhiet, 2008; Bass, 2000; Batterjee & Ashria, 2015; Costenbader & Ngari, 2001; Humble et al., 2016; Hur & Lynn, 2013; Husain et al., 2019; Kitsao-Wekulo et al., 2013; Knowles, 2008; Malda et al., 2010; Owen,

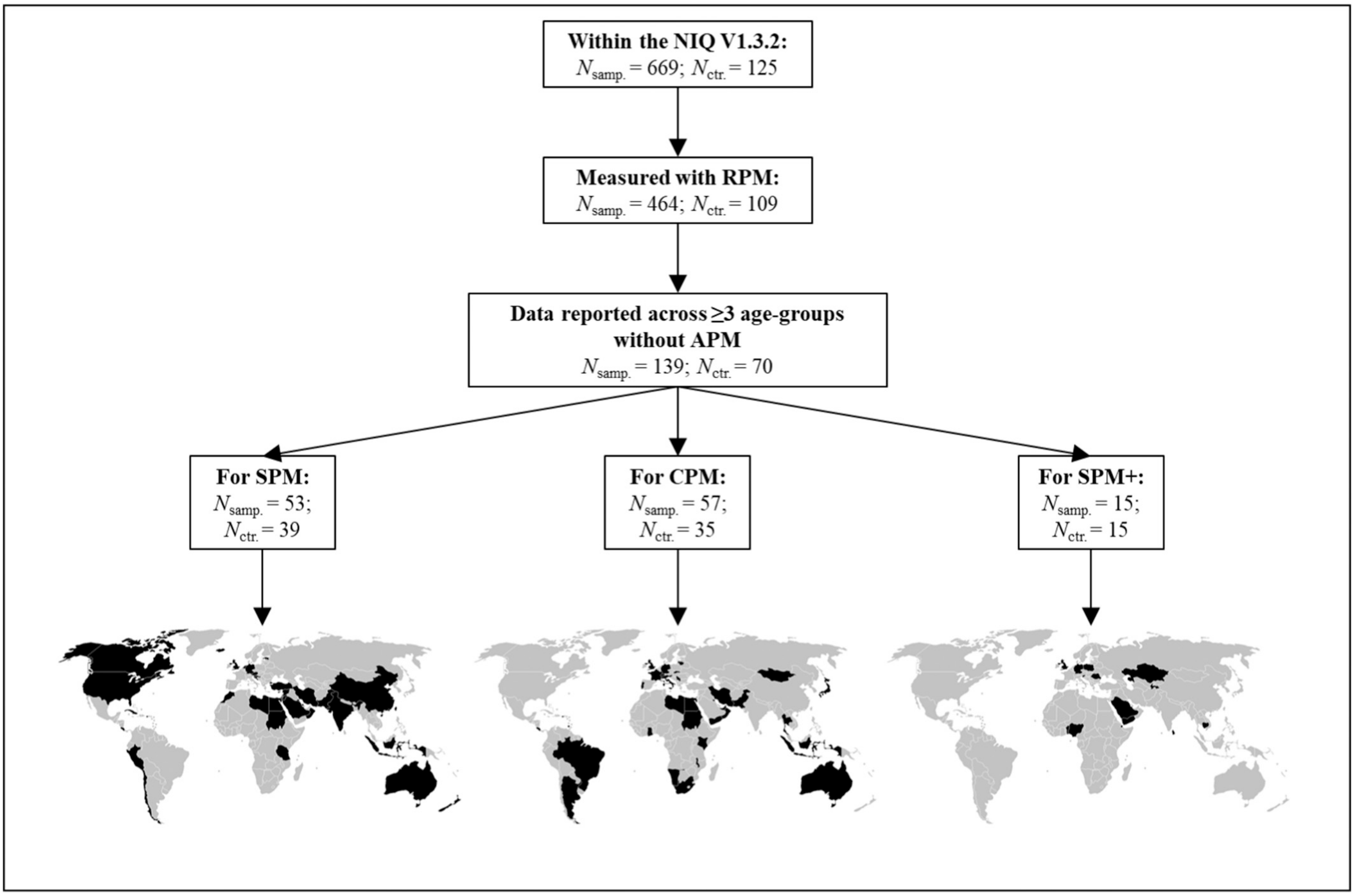

Fig. 2. Selection process of samples from the NIQ-dataset (V1.3.2) (Becker, 2019).

1992; Van de Vijver & Brouwers, 2009).

The data from Lynn and Vanhanen (2012), and Lynn and Becker (2019) has faced criticism due to occasional concerns about the quality of the samples (Hunt & Sternberg, 2006; Volken, 2003). However, there are two independent studies that measure similar theoretical constructs and provide insight into the validity of the NIQ data. Firstly, to define their variable "Learning," Lim et al. (2018) employed international school assessment data (TIMSS, PIRLS) and psychometric data (Wechsler Scales, Raven's Matrices). The correlations between their data and the NIQ data were .87 (95%$CI = .82 \mid .90$). Secondly, a correlation of .87 (95%$CI = .83 \mid .91$) exists between the NIQ data and "Harmonized Learning Outcomes (HLO)," as estimated by Angrist et al. (2021) using international and regional assessments. Furthermore, Warne (2023) conducted an analysis and concluded that, despite acknowledging a few limitations and recommending necessary enhancements to the NIQ dataset, there isn't a substantial threat to the validity of the NIQ data as estimated by Lynn and Vanhanen and Lynn and Becker.

2.2.2. Sample selection

Fig. 2 outlines the process used to select the samples utilized in this study. All these samples were drawn from the NIQ dataset (V1.3.2), which comprises 669 psychometric IQ score measurements obtained through various IQ tests on samples from 125 countries. This dataset contains raw scores gathered from administrations conducted across diverse countries, primarily focusing on age groups. This feature makes it well-suited for tracking cognitive development and age-related differences. Among the 464 samples administered in 109 nations using a variant of the RPM, not all were suitable for the study's objectives. We concentrated on age trends, necessitating the inclusion of test outcomes from a minimum of two age groups; however, we raised the threshold to three for better distinction between continuous trends and random fluctuations. Age groups were considered only if they fell within the designated age range for RPM versions: 6.5 to 15.5 years for the SPM, 4.0 to 11.5 years for the CPM, and 7.5 to 18.5 years for SPM+. Moreover, samples utilizing the APM (Advanced Progressive Matrices) were excluded, given that the APM is tailored for subjects primarily beyond the age of progressive cognitive development (14+) (Li et al., 2004). For each RPM version, a normative sample from the UK was incorporated into the dataset. This process resulted in a collection of 139 samples hailing from 70 different countries.

Out of these samples, 134 are characterized as either representative, normative, or random. This classification signifies that the test subjects were not selected based on additional criteria, such as school type, family background (including adoptive families and children's homes), disease status, treatment groups, and so forth. The absence of these additional criteria ensures that the sample reflects the population of the respective country. The remaining five samples are classified as selective and we evaluated them individually. For instance, a sample from Croatia (Žebec et al., 2015) was chosen based on right-handedness, while a sample from Poland (Dobrean et al., 2008) consisted of Army recruits. Both samples were deemed representative or, at the very least, random to a limited but acceptable extent. A similar assessment was applied to a study conducted in Malawi (Van de Vijver & Brouwers, 2009), wherein children were included in the sample only if approved by village chiefs. However, the authors argued that there was no indication of a specific mechanism employed by the village chiefs for selecting children. Two samples were selected based on the criterion of health, or the absence of a specific disease. One such sample originates from Indonesia (Bleichrodt et al., 1980) and is described as belonging to a "non-cretinous population," while the other comes from Italy (Belacchi et al., 2008) and is labeled as "healthy." The latter was retained in the study, given Italy's status as a developed Western nation, which should result in a relatively small proportion of children with disabilities that could potentially impact cognitive development. Conversely, the Indonesian sample was excluded due to the opposite assumption in the context of a developing country.

Additionally, two samples from a study conducted by Osman et al. (2017) were excluded. The data reported by the source included SPM and SPM+ scores of South Sudanese children residing in refugee camps in Sudan. As a result, these samples were deemed unrepresentative of their country of origin. The same exclusion rationale applied to the sample from Ahmed et al. (2017), which was labeled as representative but also comprised children from refugee camps in Sudan. Similarly, data from Bakhiet et al. (2017) concerning Somali refugees in Kenya were not included. Another Sudanese sample from Batterjee and Ashria (2015) was excluded due to its exclusive origin from private schools. Furthermore, three samples from Grigoriev and Lynn (2014) provided data for Kazakhs, Uzbeks, and Russians in Kazakhstan separately. Despite the fact that these three ethnicities constitute significant portions of the population in Kazakhstan, the separate age groupings for each ethnicity posed challenges in amalgamation. Consequently, the decision was made to exclude the samples for Uzbeks and Russians.

A similar approach was applied to five US samples from Raven (2000). These samples presented data for two samples of Whites, and one each of Blacks, Hispanics, and Asians. Because these samples only provided data for age groups ranging from 12.0–12.9 to 15.0–15.9, and a representative US sample spanning the entire age range of the SPM was already incorporated (Raven et al., 1999), the choice was made to exclude these five single-ethnicity samples.

Of all the intelligence measurements examined in this study, Raven's Standard Progressive Matrices (SPM) were administered across 53 samples originating from 39 distinct countries. These samples collectively span continents, encompassing regions including the Arabic-speaking world and Africa. Furthermore, Raven's Colored Progressive Matrices (CPM) were utilized for a total of 57 samples originating from 35 different countries across all continents except North America. Lastly, Raven's Standard Progressive Matrices Plus (SPM+) were employed for 15 samples hailing from 15 diverse countries across Europe, Central Asia, Sub-Saharan Africa, and the Arabian Peninsula. For a comprehensive list of the sources of psychometric IQ measurements employed in this study, please refer to Appendix I, and for an overview of these samples, consult Table A1 in Appendix II.

2.2.3. Data preparation

The data were extracted from the 'CAL' sheet of the main working file "N-IQ-DATA (V1.3.2).xlsx" within the NIQ-dataset. This sheet provides descriptive statistics for all samples for which raw scores were available for multiple age groups, as detailed in their respective sources. In cases where a sample was reported in multiple tables, such as separate tables for males and females or different age groups, those samples were merged or averaged to ensure consistency.

Normative scores from the UK and scores from the compared samples were provided for varying age groups. To facilitate comparison, these scores were standardized on a uniform age scale, as presented in Table 1. When scores were available for multiple age groups falling within the same interval on the standardized age scale (e.g., 7.3 and 7.7 both fall within the 7.0–7.9 interval), the average was calculated. Calculations for annual changes were conducted only when scores were available for two consecutive years of life. Descriptive statistics for all the data used can be found in Table A2 of Appendix II.

Raw scores were transformed into IQ scores on a scale based on British norms, referred to as "British IQ scores." This conversion was performed using a series of polynomials provided in the "NORM" sheet of the main working file "N-IQ-DATA (V1.3.2).xlsx" within the NIQ-dataset. The efficacy of this method has been successfully demonstrated by Lynn and Becker (2019, pp.26–30). Utilizing such formulas, as opposed to relying on norm tables from manuals, holds a distinct advantage, especially when converting raw scores from various groups into IQ scores. This advantage becomes apparent when dealing with decimal values that are not readily accommodated by norm tables.

Table 1
Arrangement of age groups and UK test norms across samples.

Unified age scale	SPM		CPM		SPM+	
	Age scale	RS at 50th *P*	Age scale	RS at 50th *P*	Age scale	RS at 50th *P*
			4.0–4.5	15.00		
			4.5–5.0	15.00		
5.0–5.9			5.0–5.5	17.50		
			5.5–6.0	18.00		
6.0–6.9	6.5	16.50	6.0–6.5	20.50		
	7.0	17.00	6.5–7.0	22.50		
7.0–7.9	7.5	22.50	7.0–8.0	26.50	7.0–7.9	22.50
	8.0	25.00				
8.0–8.9	8.5	31.00	8.0–9.0	28.00	8.0–8.9	26.50
	9.0	32.00				
9.0–9.9	9.5	35.00	9.0–10.0	29.50	9.0–9.9	30.50
	10.0	38.00				
10.0–10.9	10.5	39.00	10.0–11.0	31.50	10.0–10.9	32.00
	11.0	40.00				
11.0–11.9	11.5	41.00	11.0–12.0	33.00	11.0–11.9	32.50
	12.0	41.00				
12.0–12.9	12.5	42.00			12.0–12.9	33.00
	13.0	43.50				
13.0–13.9	13.5	45.50			13.0–13.9	34.50
	14.0	46.00				
14.0–14.9	14.5	48.00			14.0–14.9	34.50
	15.0	46.00				
15.0–15.9	15.5	47.00			15.0–15.9	35.50
16.0–16.9					16.0–16.9	36.00
17.0–17.9					17.0–17.9	38.00
18.0–18.9					18.0–18.9	38.00

Notes. ages given in years of life; italic norms not used; normative data given for SPM by Raven (2000), for CPM by Raven (2008a) and for SPM+ by Raven (2008b).

2.3. *Data for human development*

The multitude of factors potentially influencing the pace of cognitive development is virtually endless, spanning genetics, education, health, wealth, climate, politics, demographics, and more (Rindermann et al., 2016). Numerous studies have examined environmental factors typical of developing countries that contribute to cognitive developmental delays from early childhood through to adulthood. Such factors include the prevalence of infectious diseases such as viruses and parasites (Boivin & Giordani, 1993; Jardim-Botelho et al., 2008; Jukes et al., 2006; Leon et al., 1975; Nga et al., 2011; Nokes, Grantham-McGregor, Sawyer, Cooper, Bundy, 1992; Nokes, Grantham-McGregor, Sawyer, Cooper, Robinson, & Bundy, 1992; Peixoto & Kalei, 2013), the consequences of malnutrition, including being chronically underweight (Grantham-McGregor et al., 1994, 1997; Ijarotimi & Ijadunola, 2007), conditions such as Kwashiorkor and Marasmus (Galler et al., 1987), Konzo (Boivin et al., 2013), iodine deficiency (Bautista et al., 1982), iron deficiency (Bleichrodt et al., 1980; Soemantri, 1989), zinc deficiency (Chiplonkar & Kawade, 2014), obesity (Galván et al., 2013), exposure to toxins such as arsenic (Hamadani et al., 2011; Wasserman et al., 2007), lead (Boucher et al., 2009; Counter et al., 1998), manganese (Bouchard et al., 2011; Menezes-Filho et al., 2011), mercury (Weihe et al., 2002), and pesticides (Harari et al., 2010; Rodríguez, 2012). Additionally, inadequate socio-economic resources for upbringing and education, as well as significant disparities in accessing such resources (Dendir, 2013; Glewwe & Jaccoby, 1992; Heilmann, 2013; Iloh et al., 2017; Jukes & Grigorenko, 2010; Stein et al., 2005), neglect of school attendance due to child labor (Heady, 2003), and high rates of psychological disorders such as PTSD (de Neubourg & de Neubourg, 2011), all contribute to cognitive developmental delays.

Certain cultural practices, although hypothetical and controversial, may also play a role in some developing countries. One such example is consanguinity (Afzal, 1988; Agrawal et al., 1984; Tadmouri et al., 2009; Woodley, 2009). Regrettably, there are no meta-analyses that systematically and internationally estimate the true effects of these factors. Nonetheless, it is theoretically possible that the cumulative negative effects of any or all of these factors on children and adolescents could account for our findings. Despite the extensive research conducted on these factors, we possess limited knowledge about the specific effects of these environmental determinants. In the absence of concrete evidence, we must remain uncertain about the relative significance of these causal factors. Furthermore, it's worth considering the extent to which these factors hold importance in varying locations or cultures. Dutton et al. (2017) identified a negative Flynn effect in Kuwait, especially among older age groups, speculating that changes in the environment of school-aged children in Kuwait, such as an increased emphasis on religious indoctrination and reduced focus on science in the school curriculum, might be a contributing factor.

To account for these complexities, we incorporated the Human Development Index (HDI) by the United Nations Development Programme (UNDP, 2017) as a comprehensive indicator of living conditions. The HDI amalgamates life expectancy, education, and wealth (GNI per capita) – factors influenced by a plethora of other elements including health, nutrition, economic progress, social stability, and cultural nuances. This index exists in two forms: the standard Human Development Index (HDI) and an inequality-adjusted form (IHDI). The latter considers the equality or inequality in the distribution of human development within a society, making it more suitable but available for fewer countries. Due to the relatively limited number of cases (samples and countries) and to prevent distortions stemming from handling missing values through imputation or maximum likelihood processes (Lüdtke et al., 2007), we chose to use the IHDI score if available, otherwise opting for the HDI. This amalgamated variable is referred to as the Combined Human Development Index (CHDI) in subsequent discussions. This decision is justified by the high correlations observed between the two HDI forms ($r = .99$ across 152 countries according to the NIQ-dataset 1.3.2). It is important to note that this process was carried out without using regression equations, but by equalizing means and standard deviations for countries that had both types of data available.

2.4. *Analyses*

2.4.1. *Within-sample analyses*

In our initial analysis, we seek to address the question: Does intelligence exhibit an age-related divergence or convergence, causing the intelligence of a sample to deviate from the UK norms as individuals' age increases? The ideal approach would involve calculating effect sizes of maturation across different samples and then comparing these with the effect size of the UK sample. Unfortunately, neither the NIQ-dataset nor the original sources provide standard deviations for all samples. For instance, even across all 22 samples from Raven's manuals, including the UK norm samples, standard deviations or variances are not provided. Alternatively, we could have tackled each sample separately, calculating the differences in raw scores for each of its age groups compared to the corresponding age-group in the UK norms, and then correlating these differences with age. Although such a correlation coefficient wouldn't offer insight into the magnitude of the Simber Effect, it would effectively capture its direction and clarity. However, two significant issues arise: First, a difference of x raw scores in one age group doesn't represent the same real-world intelligence difference as a difference of x raw scores in another age group. Thus, it is more appropriate to first convert raw scores into British IQ scores and then proceed with the aforementioned procedure, but now using differences in IQ scores (IQD).

A second issue emerges due to the possibility of positive or negative differences in IQ scores. To address this, we employ the absolute difference ($|\mathrm{IQD}|$). Furthermore, it's essential to recognize that we cannot distinguish whether a negative correlation coefficient signifies that an IQ score below the British 100 is increasing its deviation as age advances or that an IQ score above 100 is approaching 100 as age increases. Ultimately, a positive correlation between $|\mathrm{IQD}|$ and age indicates an increase in IQ score differences from the UK as age progresses, whether positively or negatively, and conversely, a negative correlation between |IQD| and age suggests a decrease in IQ score differences from the UK by age, in either a positive or negative manner.

2.4.2. Across-sample analyses

If the analysis described in Section 2.4.1 yields evidence of any type of Simber Effect, subsequent analyses will investigate whether the Simber Effect is random or patterned – specifically, whether it's related to living conditions or human development. To assess this, we will compare the mean absolute differences in human development from the UK (|CHDID|) for countries whose sample data show increasing divergence in intelligence with those of the entire set of samples. This comparison will help determine if countries with increasing divergence are more dissimilar in terms of human development compared to the entire sample set.

However, due to the limitations of the previous analyses in determining the strength of potential Simber Effects, cross-sample regression analyses based on their results are currently not viable. A significant challenge arises from the diverse composition of age groups in the samples. Since we can't assume that Simber Effects will remain constant across the entire lifespan, and potential shifts from divergence to convergence (or vice versa) need consideration, the varying mean ages of the samples could significantly impact the outcomes of regression analyses. We intend to address this challenge using two approaches:

First, individual annual shifts between age groups will be considered – for instance, the shift from 4.0 to 4.9 to 5.0–5.9 (4 → 5) or from 8.0 to 8.9 to 9.0–9.9 (8 → 9). Changes in raw scores for each shift in each sample (RS↕) will be calculated where data are available. Subsequently, RS↕ will be correlated across all samples with NIQ and CHDI separately for each shift. In this analysis, we can use raw scores instead of IQ scores, as the use of uniform age groups within each correlation eliminates the need for age normalization. Despite necessitating separate consideration for the three observed Raven's tests due to their distinct raw score scales, such differentiation was already planned to explore the impact of different test usage on patterns discovered. Additionally, this approach enhances clarity by allowing an increase in raw scores to correspond with progressive cognitive development, whereas even a decrease in IQ scores can indicate progressive cognitive development, albeit at a slower pace than the UK.

Second, for multivariate regression analysis, the samples' mean RS↕ will serve as the dependent variable, with NIQ, CHDI, mean sample age (Age), and the year the RPM was administered (YoA) as independent variables. This method aims to quantify the portion of explained variance in cognitive development speed attributed to environmental conditions rather than age differences between samples, average intelligence level, or the Flynn Effect. Since multicollinearity could be an issue due to the likely positive correlation between NIQ and CHDI, separate models will be run for SPM, CPM, and SPM+ − one with NIQ and one without NIQ. These multivariate regression analyses were performed using MPlus (©Muthen & Muthen, 1998–2015) with the full information maximum likelihood (FIML) to handle missing values. It is worth noting that MPlus might encounter unexpected computation problems if variances in variables are excessively large in terms of absolute values. Consequently, all variables were z-standardized before being used in multivariate analyses to mitigate this issue.

2.4.3. Simulation for balanced out Simber Effects

If systematic Simber Effects are indeed present, the question arises as to how IQ scores would appear if the RS↕ values observed in the samples were replaced by RS↕ values from the UK. To perform such a simulation, the raw score for the youngest available age group within a sample should be selected and modified each year according to the corresponding change observed in the UK. This process will yield two sets of age-specific raw score trends, indicating different levels of intelligence but with a perfect correlation between them. An illustrative simulation is depicted in Fig. 3. Since the youngest available age group varies among the samples (resulting in weaker correction for a sample with an older lowest age group compared to a sample with a younger lowest age group), this method cannot fully compensate for the complete variances in RS↕, yet it might account for a significant portion.

Ultimately, the measured and simulated raw scores need to be converted to IQ scores using the procedure described in Section 2.4.1. Additionally, the conversion should factor in norm inflation (using the "FEC" sheet of the NIQ-dataset main working file "N-IQ-DATA (V1.3.2).xlsx"; see Lynn & Becker, 2019, p. 32) before conducting a comparison.

2.5. Annotations

We will provide or indicate *p*-values for all analyses, although achieving the lowest levels of statistical significance is unlikely due to the relatively small number of cases in each of the 8 to 12 age groups, depending on the specific RPM, or 15 to 57 in cross-sample analyses. Nevertheless, the stability of observed patterns across multiple analyses conducted with different samples serves as a more informative criterion.

It is important to reiterate that the collected samples are solely obtained through cross-sectional designs. Hence, when referring to "aging" for simplicity, we are not necessarily implying a genuine longitudinal maturation process in the context of gerontology. Rather, we are comparing two different age groups. We will delve into this distinction in greater detail later.

For a more comprehensive understanding, Appendix III includes graphical representations of all selected samples.

3. Results

3.1. Within-sample analyses

Tables 2, 3, and 4 present the correlations between age and the absolute difference in IQ scores between the sample and the UK norms ($r_{\text{Age}\leftrightarrow|\text{IQD}|}$). For 13 samples, no correlations were computed as data were reported for only three separate age groups. Among the remaining 109 samples, 90 (83 %) exhibit correlations with $|r| \geq .30$, indicating either an increase or decrease in IQ score differences compared to the UK as age advances. A minority of 22 (20 %) of these samples also attain statistical significance with $p \leq .05$. Among those with $|r| \geq .30$, the positive/negative ratio is 68:22, revealing that more samples display an augmented difference to the UK norms with increasing age. This pattern persists when examining RPM groups individually: 76 % of the SPM samples, 90 % of the CPM samples, and 75 % of the SPM+ samples exhibit correlations with $|r| \geq .30$. The ratios are 22:13, 39:7, and 7:2, respectively. Thus, a substantial majority of the samples indicate increasing divergences.

3.2. Across-sample analyses

The mean |CHDID| for all 103 samples (6 were excluded due to missing HDI data) is .1861 (N = 103; SD = .1592). Among the 64 samples with $r_{\text{Age}\leftrightarrow|\text{IQD}|} \geq .30$, the mean |CHDID| is .2142 (N = 64; SD = .1597). Within the SPM group, it is .2056 (N = 20; SD = .1479), within the CPM group it is .2106 (N = 37; SD = .1581), and within the SPM+ group, it is .2580 (N = 7; SD = .1928). The respective *t*-values are −1.108 (p = .269) for all groups, −.507 (p = .613) for the SPM group, −.804 (p = .423) for the CPM group, and − 1.142 (p = .256) for the SPM+ group. Although samples with increasing divergence show larger differences on average in terms of human development compared to the UK, these differences are not statistically significant; however, they remain quite robust.

It is important to acknowledge the potential impact of country unrepresentativeness within the different RPM groups (e.g., seven samples from the SPM group are from Sudan). By averaging $r_{\text{Age}\leftrightarrow|\text{IQD}|}$ (via Fisher transformation) and |CHDID| across 103 samples to obtain 59 country

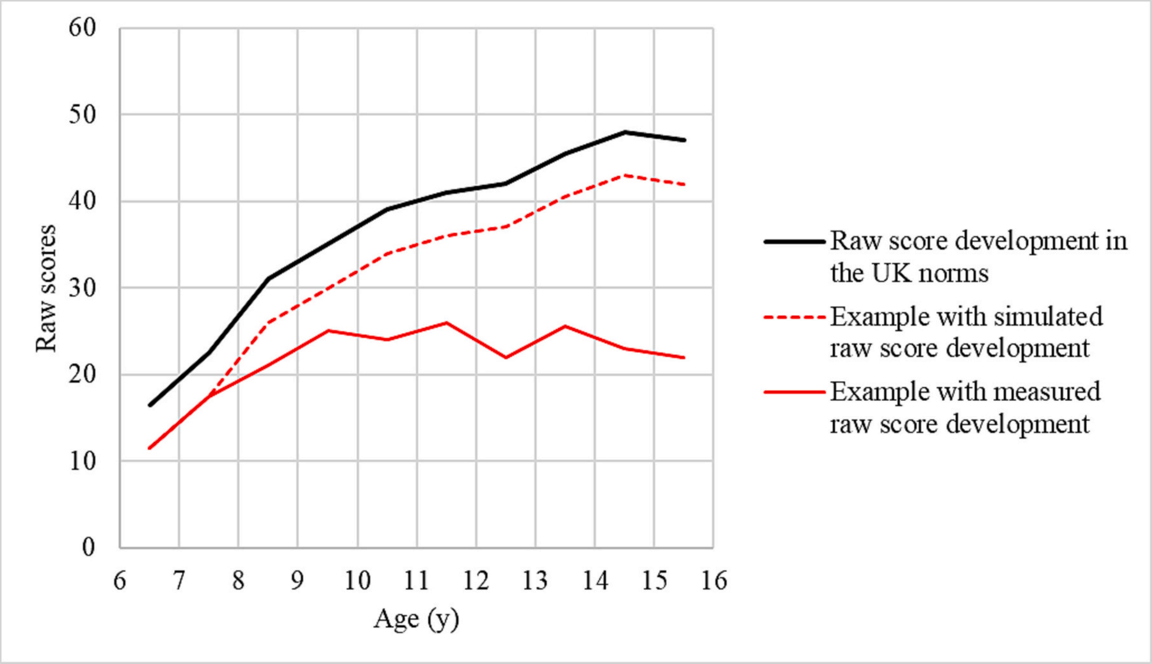

Fig. 3. Example for simulated (out-balanced) raw score development.

Table 2
Correlations between age and absolute IQD within each sample of the SPM group.

ID	N_{AG}	$r_{\text{Age}\leftrightarrow\|\text{IQD}\|}$	p	ID	N_{AG}	$r_{\text{Age}\leftrightarrow\|\text{IQD}\|}$	p
AUS4319	8	.44	.273	NZL2045	8	.43	.289
AUT4153	3	n.c.	n.c.	OMN7680	6	−.40	.432
CAN8614	3	n.c.	n.c.	PAK2866	4	.35	.649
CHL1702	5	.36	.554	PER5808	3	n.c.	n.c.
CHN9787	10	−.37	.298	PRI6308	8	−.38	.357
CRI7021	4	.90	.099	PSE1534	8	.31	.457
CUB6365	4	.77	.231	QAT1366	6	.64	.173
CYP6250	6	−.57	.240	SAU1749	8	.69	.058
DEU7126	5	−.86	.064	SAU6541	8	.82	.012
EGY9492	10	−.77	.009	SAU8865	8	.19	.649
EST8199	9	−.25	.510	SDN1289	7	−.71	.076
HKG6221a	8	−.27	.525	SDN1388a	5	.90	.038
HKG6221b	10	.77	.009	SDN1388b	5	.84	.077
HRV2479	6	.56	.253	SDN2829	7	.38	.403
IDN6764	4	.09	.912	SDN5606	5	.38	.530
IND5561	5	.21	.729	SDN7769	7	−.41	.356
IND7068	8	−.65	.079	SDN8415	7	.40	.371
IRN8850	7	.24	.603	SVN2723	8	−.20	.643
ISL2755	10	−.50	.140	SYR4283	3	n.c.	n.c.
ISR7998	6	−.72	.107	SYR4825	3	n.c.	n.c.
JOR2837	7	−.15	.750	SYR6869	8	.59	.127
KWT5643	8	−.85	.007	TUR4222	9	.73	.025
KWT9995	8	−.65	.081	TWN6602	4	.45	.548
LBY7870	8	.25	.544	TWN7259	7	.61	.148
MAR4071	4	.81	.187	TZA8412	3	n.c.	n.c.
NLD5358	7	−.26	.578	USA5299	10	−.07	.850

Notes. Positive *r* indicate bigger differences in higher ages; IDs from NIQ-dataset (V1.3.2) (Becker, 2019), marked with a and b if more than one sample with the same ID; |IQD| = absolute IQ differences to UK norm sample; n.c. = not calculated (if $N \leq 3$).

means (refer to Table A4 of Appendix II for data), 37 countries exhibit $r_{\text{Age}\leftrightarrow|\text{IQD}|} \geq .30$, with a mean |CHDID| of .2180 ($N = 37$, $SD = .1647$), compared to .1813 ($N = 59$, $SD = .1533$) for all countries. The calculated *t*-value is −1.109 ($p = .270$), again indicating insignificance but robustness. Collectively, this evidence suggests that the cross-national divergence in IQ scores is not random but rather patterned.

Table 3
Correlations between age and absolute IQD within each sample of the CPM group.

ID	N_{AG}	$r_{\text{Age}\leftrightarrow\|\text{IQD}\|}$	p	ID	N_{AG}	$r_{\text{Age}\leftrightarrow\|\text{IQD}\|}$	p
ARE2499	6	.63	.177	LBY9541	6	.91	.011
ARG8081	6	−.04	.944	LBY9633	3	n.c.	n.c.
AUS4314	6	.49	.320	LTU5971	6	.73	.102
AUS5624	6	−.57	.238	MNG3257	7	.38	.402
BRA4066	7	.72	.067	MWI4457	5	.94	.017
BRA4770	7	.83	.021	NAM3261	4	.89	.108
BRA7075a	4	.44	.562	OMN1923	7	.88	.008
BRA7075b	4	.95	.049	OMN4512	3	n.c.	n.c.
CHE4251	8	.21	.615	OMN6270	7	.87	.011
CHE6518	5	−.29	.634	PAK6352	3	n.c.	n.c.
CHE7313	5	−.55	.332	PRI5582	7	.52	.227
CRI3498	3	n.c.	n.c.	PRI6157	7	.67	.098
DEU6886	7	.78	.041	PRT4450	6	.21	.689
EGY7781	7	.87	.011	PSE5248a	6	.76	.081
EGY9075	5	.35	.559	PSE5248b	7	.45	.315
FRA2835	8	−.46	.247	SDN3606	4	.97	.030
GHA7539	5	.97	.007	SDN5195	4	.98	.025
IDN3746	4	.72	.278	SRB6215	8	.89	.003
IRN3616	5	.45	.447	SVK2338	6	−.42	.412
ITA2558	5	−.49	.402	SVN5820	6	.86	.028
ITA4253	8	.57	.137	SVN9451	6	.83	.040
ITA4819	7	−.07	.889	THA3965	5	.68	.209
ITA7980	5	.39	.518	THA7616	5	−.85	.066
JPN1123	4	−.94	.055	THA9367	5	.88	.051
KEN6039	5	.76	.140	VCT6406	4	.89	.107
KWT1478	5	.59	.293	YEM8656	4	.75	.250
LBY3196	6	.75	.088	ZAF7089	3	n.c.	n.c.
LBY5613	6	.72	.105	ZAF7824	6	.78	.070

Notes. Positive *r* indicate bigger differences in higher ages; IDs from NIQ-dataset (V1.3.2) (Becker, 2019), marked with a and b if more than one sample with the same ID; |IQD| = absolute IQ differences to UK norm sample; n.c. = not calculated (if $N \leq 3$).

Regarding SPM and CPM (as shown in Table 5), the correlation between RS↕ and NIQ or CHDI is mostly positive, particularly stronger in the younger age groups than in the older ones. The association direction reverses for SPM between 9 and 12 years of age, and for CPM, the

Table 4
Correlations between age and absolute IQD within each sample of the SPM+ group.

ID	N_{AG}	$r_{Age\leftrightarrow\lvert IQD\rvert}$	p	ID	N_{AG}	$r_{Age\leftrightarrow\lvert IQD\rvert}$	p
BEN3861	12	.54	.070	POL5086	3	n.c.	n.c.
DEU2018	6	−.23	.655	ROU4821	12	−.69	.012
DJI3851	10	.55	.102	SAU9747	6	−.93	.007
KAZ5019	8	.42	.300	SDN4915	12	−.24	.447
KHM4764	12	−.08	.812	SVN3452	7	.58	.175
LKA6243	6	.41	.419	TJK5305	3	n.c.	n.c.
NGA8922	10	.59	.071	YEM7411	7	.68	.090

Notes. Positive r indicate bigger differences at higher ages; IDs from NIQ-dataset (V1.3.2) (Becker, 2019), marked with a and b if more than one sample with the same ID; |IQD| = absolute IQ differences to UK norm sample; n.c. = not calculated (if $N \le 3$).

Table 5
Correlations of RS↕ with NIQ and CHDI across countries for each shift between two age groups separately.

Shift between age groups	SPM		CPM		SPM+	
	NIQ	CHDI	NIQ	CHDI	NIQ	CHDI
4.0–4.9 to 5.0–5.9			−.59	−.62		
5.0–5.9 to 6.0–6.9			.37	.06		
6.0–6.9 to 7.0–7.9	.64	.69*	.51*	.52*		
7.0–7.9 to 8.0–8.9	.66*	.63*	.40*	.34*	.75*	.67
8.0–8.9 to 9.0–9.9	.51*	.42*	.30*	.36*	−.02	.12
9.0–9.9 to 10.0–10.9	.04	.09	.00	.03	.12	.32
10.0–10.9 to 11.0–11.9	−.12	−.14	.00	−.05	−.05	.12
11.0–11.9 to 12.0–12.9	−.04	.00			.23	.16
12.0–12.9 to 13.0–13.9	−.20	−.17			.07	−.09
13.0–13.9 to 14.0–14.9	−.32	−.38*			−.08	−.11
14.0–14.9 to 15.0–15.9	−.08	.04			.37	.00
15.0–15.9 to 16.0–16.9					.36	.22
16.0–16.9 to 17.0–17.9					.21	.50
17.0–17.9 to 18.0–18.9					.43	.51

Notes. $N_{spl.}$ varies between 16 and 40 for SPM, 6 and 52 for CPM, 7 and 12 for SPM+.

* $p \le .05$.

correlation vanishes at the same age. There is only one exception: the negative correlations in the youngest age groups for CPM. When combining SPM and CPM data, it appears that the association between RS↕ and NIQ or CHDI is positive within the age range of 6 to 9 years—typically the initial three years of schooling—while being minimal in preschoolers and children older than 9 years. In the case of SPM+, no distinct pattern emerged, possibly due to the limited number of samples.

Taken together, these findings suggest that samples from countries with higher NIQ and CHDI experience more pronounced increases in raw scores during the younger age groups, while exhibiting weaker increases or even decreases in raw scores during the older age groups. It is essential to note that not all samples provided raw scores for each age group, which could influence the correlations presented in Table 5. However, it is noteworthy that the observed pattern remains consistent in both the SPM and CPM groups, and to some extent, in the SPM+ group as well.

Table 6
Results from two multivariate analysis in two models predicting annual raw score changes for SPM.

	Factors	r	β	$S.E.$	p
Model 1	NIQ	−.01	−.12	.196	.531
	CHDI	.03	.00	.176	.980
	Age	−.40	−.40	.132	.002
	YoA	.10	.03	.150	.860
	Model parameter	Value		$S.E.$	p
	$logL(\theta)$	−316.612		–	–
	ν	−.005		.132	.969
	$1 - R^2$	.847		.095	.000

	Factors	r	β	$S.E.$	p
Model 2	CHDI	.03	−.06	.141	.659
	Age	−.37	−.38	.129	.003
	YoA	.10	.07	.136	.624
	Model parameter	Value		$S.E.$	p
	$logL(\theta)$	−269.741		–	–
	ν	.001		.133	.995
	$1 - R^2$	.854		.093	<.001

Notes. dependent variable = RS↕; $N_{spl.}$ = 49; used values z-standardized; estimator = FIML; NIQ = national IQ, CHDI = combined Human Development Index, YoA = year of test administration.

Table 7
Results from two multivariate analysis in two models predicting annual raw score changes for CPM.

	Factors	r	β	$S.E.$	p
Model 1	NIQ	.46	.05	.197	.792
	CHDI	.47	.41	.178	.021
	Age	−.08	−.15	.124	.237
	YoA	−.28	−.22	.127	.085
	Model parameter	Value		$S.E.$	p
	$logL(\theta)$	−346.896		–	–
	ν	−.001		.114	.995
	$1 - R^2$	.710		.103	<.001

	Factors	r	β	$S.E.$	p
Model 2	CHDI	.47	.45	.105	<.001
	Age	−.08	−.16	.114	.163
	YoA	−.28	−.23	.114	.040
	Model parameter	Value		$S.E.$	P
	$logL(\theta)$	−298.879		–	–
	ν	.001		.114	.994
	$1 - R^2$	.711		.103	<.001

Notes. dependent variable = RS↕; $N_{spl.}$ = 55; used values z-standardized; estimator = FIML; NIQ = national IQ, CHDI = combined Human Development Index, YoA = year of test administration.

The utilization of the samples' mean ARS↕ along with NIQ, CHDI, age, and YoA in multivariate analyses (as shown in Tables 6 to 8) reveals that age serves as the most effective predictor of RS↕ for the SPM group in model 1 ($\beta = -.40$; $r = -.40$; $S.E. = .132$; $p = .002$) and model 2 ($\beta = -.38$; $r = -.37$; S.E. = .129; $p = .003$). This phenomenon can be attributed to a ceiling effect of the test at older ages. For the CPM group, the combined human development index (CHDI) emerges as the strongest predictor of annual score gains in model 1 ($\beta = .41$; $r = .47$; $S.E. = .178$; $p = .021$) and model 2 ($\beta = .45$; $r = .47$; $S.E. = .105$; $p < .001$). This indicates an accelerated cognitive maturation in countries with higher levels of human development. No statistically significant predictors of annual score gains are identified for the SPM+ group, possibly due to the limited number of available samples.

3.3. Simulation for balanced out Simber Effects

Assuming uniform cognitive development speed across all samples, in contrast to the actual measured rates, resulted in noticeable increases in the IQ scores of the samples, with the exception of the SPM+ group (refer to Table 9, and for detailed data, see Table A3 in the Appendix II). However, the correlation between real and simulated IQ scores is remarkably high at both the cross-sample level ($r = .87$; $N_{smpl.} = 125$; $p < .001$) and when aggregated by countries ($r = .91$; $N_{ctry.} = 68$; $p < .001$).

The degree of change in IQ scores through simulation exhibited correlations with both NIQ ($r_{RPM} = -.27$; $r_{SPM} = -.07$; $r_{CPM} = -.38$; $r_{SPM+} = -.43$) and CHDI ($r_{RPM} = -.25$; $r_{SPM} = -.07$; $r_{CPM} = -.36$; r_{SPM+}

Table 8
Results from two multivariate analysis in two models predicting annual raw score changes for SPM+.

	Factors	*r*	*β*	*S.E.*	*p*
Model 1	NIQ	.19	−.07	.378	.857
	CHDI	.35	.25	.398	.530
	Age	−.08	−.27	.301	.368
	YoA	−.28	−.30	.372	.413
	Model parameter	Value		*S.E.*	*p*
	$logL(\theta)$	−87.680		–	–
	ν	.000		.234	1.000
	$1 - R^2$	.821		.179	<.001

	Factors	*r*	*β*	*S.E.*	*p*
Model 2	CHDI	.35	.21	.317	.516
	Age	−.08	−.29	.285	.311
	YoA	−.28	−.30	.114	.420
	Model parameter	Value		*S.E.*	*p*
	$logL(\theta)$	−74.157		–	–
	ν	.000		.234	1.000
	$1 - R^2$	.822		.179	<.001

Notes. dependent variable = RS↕; $N_{spl.} = 15$; used values z-standardized; estimator = FIML; NIQ = national IQ, CHDI = combined Human Development Index, YoA = year of test administration.

Table 9
Compared sample IQs before and after correction for Simber effect (SEC).

Group	*N*	With SEC		Without SEC		*r*	*d*
		M	*SD*	*M*	*SD*		
RPM	125	84.00	13.68	87.04	14.02	.87	−.22
SPM	53	84.46	11.90	85.63	12.59	.93	−.10
CPM	57	83.49	14.44	89.08	15.01	.87	−.38
SPM+	15	84.33	17.21	84.27	14.83	.86	.00

= −.48), indicating that countries with lower IQ scores or lower levels of development in comparison to the UK experienced greater corrections in their IQ scores. This alignment with our findings from Section 3.2 reinforces our conclusions.

4. Discussion

4.1. *Primary findings*

Our study yielded indications that the Simber Effect, as defined by Bakhiet et al. (2018), extends beyond the Arab world and is particularly prominent in samples from developing countries. However, this observation is somewhat constrained, as it was only detected in case of one out of the three observed Raven's tests, namely the SPM. It's worth noting that our analysis was based on a limited set of 15 samples that used the SPM+, which limits the weight of the findings for this specific test. Furthermore, the applicability of the CPM is constrained to children up to 11 or 12 years of age, potentially placing it outside the age range where the Simber Effect may occur. Given these considerations, the significant effect found for the SPM carries more significance than the non-detection in the SPM+ and CPM applications.

Our findings indicated that cognitive development occurs at a delayed pace in countries with lower levels of human development. As a result, the differences in Raven's raw scores and IQ scores between these countries and more developed and affluent nations tend to become more pronounced as the compared individuals age, up until around 9 years old. However, an intriguing question arises: why does this trend halt at this point? One possible explanation is the lower quality of primary school education in less developed countries. Yet, this explanation doesn't fully address why the trend doesn't persist into older ages. Our hypothesis is that this is due, in part, to the amalgamation of results from various school types within many studies. For instance, a study by Husain et al. (2019) in Sudan includes pupils from primary schools for younger ages and those from secondary schools or even colleges for older ages. In many studies, the samples are primarily composed of children who are attending school, and there is minimal representation of those who are not enrolled. While this is not as problematic for CPM studies, since they mainly involve primary school children, it becomes a substantial issue for SPM and SPM+ studies that encompass older age groups recruited from high schools. Individuals who struggled in primary school are more likely to drop out of the education system. In many less developed countries, not all students progress from primary to secondary school. This selectivity effect could potentially mask the Simber Effect at ages when transitions between different school levels occur, typically around ages 10 and above. Interestingly, our findings align with this pattern.

However, it is crucial to underline that the Simber Effect we identified in less developed countries through this study is subtle and not consistently robust. Even though a trend emerged suggesting that cognitive development occurs at slower rates in less developed countries compared to more developed ones, this pattern was neither strongly pronounced nor capable of explaining a significant proportion of the variance in development speeds. Yet, these observations should not hastily lead to the conclusion that the genuine effect is inherently unsystematic and of minimal significance. The irregularities and uncertainties we encountered could be attributed to the challenges associated with comparing diverse samples and the relatively low number of samples available per country. Additionally, it is important to recognize that while developing countries often exhibit lower school attendance rates, the samples from these countries are typically drawn from school populations, making them less representative of the entire national population compared to samples from Western countries (Husain et al., 2019). Consequently, a considerable portion of the actual developmental delay might be absent in these samples, as key negative factors like inadequate schooling may not be present.

Moreover, the origins of the effect remain unclear; whether it is a result of intelligence development during maturation or merely age-related disparities stemming from different living conditions during specific childhood phases across varying age groups. The latter possibility could account for the observed inconsistencies in patterns, given that the fluid and frequently changing living conditions in developing countries, often marked by crises, can lead to rapid shifts in environmental variables. However, such fluctuations might exert a lasting impact such that impoverished conditions early in development lead to lower annual developmental strides even in later life stages. Consequently, children affected by adverse circumstances not only experience an immediate decline in IQ scores but also endure long-term repercussions in the form of sluggish developmental rates, resulting in cumulative deficits throughout their lives.

Therefore, one potential explanation is that a segment of the Simber Effect observed in cross-sectional studies is an outcome of ongoing Flynn Effects in less developed countries. This would be the case when the Flynn Effect, defined as a cohort trend involving rising cognitive test scores, is more prominent in a developing nation than in the UK. Indications exist that the Flynn Effect has plateaued in numerous European countries, while it remains in progress, albeit at varying rates, in most developing countries due to ongoing economic advancement and enhancements in educational systems (Meisenberg, 2014). Consequently, older adolescents could potentially exhibit relatively lower scores than young children in cross-sectional studies simply due to their representation of earlier birth cohorts. This aspect might significantly contribute to the Simber Effect (or even mimic it) in countries that have witnessed rapid improvements in environmental conditions and educational quality in recent years.

It's important to consider that aside from environmental conditions, genetic factors could also play a role in influencing cognitive development speed. Bakhiet et al. (2018) speculated that there might be a genetic dimension to the Simber Effect, potentially linked to a fast life history strategy. Controversially, Rushton (1995) suggested genetic racial differences in life history strategy, positing adaptations to varying ancestral ecologies characterized by different degrees of stability and harshness. Further research expanded this perspective beyond Rushton's "three major races" (Sub-Saharan Africans, Caucasians, and East Asians) to include multiple of the classical anthropology's 12 races (as seen in Dutton, 2018). Intelligence, geographically, correlates with a slow life history strategy associated with adaptation to a harsh yet stable ecological environment. These patterns could imply, aligning with Bakhiet et al.'s proposal that a portion of the Simber Effect might originate from genetic factors, reflecting population disparities in life history strategy. In simpler terms, an aspect of the Simber Effect might stem from certain populations adapted to an ancestral ecology that's unstable yet not harsh, leading them, relatively, to mature cognitively at younger ages compared to slow life history strategy populations. However, their cognitive development peaks at a lower level, giving rise to the Simber Effect. This, of course, should be treated as a highly speculative notion.

A potential direction for testing this genetic hypothesis in the future could involve examining whether countries with a pronounced Simber Effect tend to experience puberty at earlier ages compared to countries without such an effect. Notably, in Western countries, puberty onset has been decreasing since the 19th century due to changes in nutrition and other environmental factors. A meaningful comparison might involve assessing the age of puberty in a wealthy country like Saudi Arabia today versus Europe today, and comparing the age of puberty in a less affluent nation like Sudan with Europe around 1900, a time when European living standards resembled those of present-day Sudan. The overarching Life History Theory suggests the existence of a "*g* factor" for life history influenced by both genes and the environment, affecting both physical and psychological maturation.

Irrespective of these genetic speculations, it is crucial to underscore the significance of accounting for age differences in comparisons of country IQ scores (and potentially other psychological metrics). Most cognitive test studies are conducted within national educational systems. While results obtained during primary school age may offer a representative snapshot for that particular age group in countries with universal primary school attendance, caution is necessary when extrapolating these findings to young adult IQ scores, especially when comparing with more developed countries like the UK. We have demonstrated the presence of the Simber Effect in early and mid-primary school years (ages 6–9), and the absence of a consistent Simber Effect in ages 10 and above might stem from increased selectivity in school attendance as children grow older in most developing countries. Consequently, disparities in average IQ scores among older adolescents and young adults between countries are likely underestimated when these estimates are derived from studies involving younger children.

Considering that most applied research pertains to country-level outcomes involving adult IQ scores (e.g., economic productivity, political stability, technological innovation), future endeavors should focus on adjusting estimates of young adult IQ scores for the Simber Effect (and ongoing Flynn Effects) whenever country-level IQ estimates are derived from studies involving younger children:

Ethical approval

No data were collected from human participants during the research. The documents were examined in this study. All ethical standards were taken into consideration and followed during the research.

Declaration of competing interest

No potential conflict of interest was reported by the author(s).

Data availability

No data was used for the research described in the article.

Acknowledgement

The authors extend their appreciation to the Researchers Supporting Project number (RSPD2023R705), King Saud University, Riyadh, Saudi Arabia, for funding this work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.actpsy.2023.104015.

References

Afzal, M. (1988). Consequences of consanguinity on cognitive behavior. *Behavior Genetics, 18*, 583–594.

Agrawal, N., Sinha, S. N., & Jensen, A. R. (1984). Effects of inbreeding on Raven matrices. *Behavior Genetics, 14*, 579–585.

Ahmed, A. S. F., Osman, H. A., Bakhiet, S. F. A., Cheng, H., & Lynn, R. (2017). Gender differences on the standard progressive matrices of south Sudanese living in Khartoum. *Mankind Quarterly, 58*, 2.

Anderson, H. E., Kern, F. E., & Cook, C. (1968). Sex, brain damage, and race effects in the progressive matrices with retarded population. *Journal of Social Psychology, 76*, 207–211.

Angrist, N., Djankov, S., Goldberg, P. K., & Patrinos, H. A. (2021). Measuring human capital using global learning data. *Nature, 592*, 403–408. https://doi.org/10.1038/s41586-021-03323-7

Bakhiet, S., Dutton, E., Ashaer, K., Essa, Y., Blahmar, T., Hakami, S., & Madison, G. (2018). Understanding the Simber effect: Why is the age-dependent increase in children's cognitive ability smaller in Arab countries than in Britain? *Personality and Individual Differences, 122*, 38–42. https://doi.org/10.1016/j.paid.2017.10.002

Bakhiet, S. F. A. (2008). Application of progressive matrices in the Sudan. *Arab Psynet, 19*, 223–225.

Bakhiet, S. F. A., Abdelrasheed, N. S. G., Al-Khadher, M. M. A., Adan, M. H., Cheng, H., & Lynn, R. (2017). A study of sex differences in intelligence of Somali refugees in Kenya. *Mankind Quarterly, 58*, 301–306.

Barbey, A. K., Colom, R., Solomon, J., Krueger, F., Forbes, C., & Grafman, J. (2012). An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain, 135*(2), 1154–1164. https://doi.org/10.1093/brain/aws021

Bass, N. (2000). *The Raven's Coloured progressive matrices test: A pilot study for the establishment of normative data for Xhosa speaking primary school pupils in the Grahamstown region*. Grahamstown, South Africa: MA Social Science, Rhodes University.

Batterjee, A. A., & Ashria, I. (2015). Intelligence and education: The Sudan case. *Mankind Quarterly, 56*, 135–166.

Bautista, A., Barker, P. A., Dunn, J. T., Sanchez, M., & Kaiser, D. K. (1982). The effects of oral iodized oil on intelligence, thyroid status, and somatic growth in school-age children from an area of endemic goiter. *The American Journal of Clinical Nutrition, 35*, 127–134.

Becker, D. (2019). The NIQ-dataset (V1.3.2). http://viewoniq.org/?page_id=9.

Becker, D., Meisenberg, G., Dutton, E., Bakhiet, S. F., Humad, O. A. M., Abdoulaye, H. A., & Ahmed, S. A. E. S. (2022). Factor structure in Raven's progressive matrices plus in sub-Saharan Africa – Benin and Djibouti. *Journal of Psychology in Africa, 32*(2), 103–114. https://doi.org/10.1080/14330237.2022.2028080

Belacchi, C., Scalisi, T. G., Cannoni, E., & Cornoldi, C. (2008). *Analisi della distribuzione descrittiva ai punteggi CPM del campione normativo, per fascia di età e per tipo di somministrazione (N = 4678)*. Firenze: Giunti O.S. Organizzaazioni Speciali.

Bleichrodt, N., Drenth, P. J. D., & Querido, A. (1980). Effects of iron deficiency on motor and psychomotor abilities. *American Journal of Physical Anthropology, 53*, 55–67.

Boivin, M. J., & Giordani, B. (1993). Improvements in cognitive performance for schoolchildren in Zaire following an iron supplement and treatment for intestinal parasites. *Journal of Pediatric Psychology, 18*, 249–264.

Boivin, M. J., Okitundu, D., Bumoko, G. M.-M., Sombo, M. T., Mumba, D., Tylleskar, T., … Tshala-Katumbay, D. (2013). Neuropsychological effects of konzo: A neuromotor disease associated poorly processed cassava. *Pediatrics, 131*, 1231–1239. https://doi.org/10.1542/peds.2012-3011

Bouchard, M. F., Sauvé, S., Barbeau, B., Legrand, M., Brodeur, M.-E., Bouffard, T., … Mergler, D. (2011). Intellectual impairment in school-age children exposed to manganese from drinking water. *Environmental Health Perspectives, 119*, 138–143. https://doi.org/10.1289/ehp.1002321

Boucher, O., Muckle, G., Saint-Amour, D., Dewailly, É., Ayotte, P., Jacobson, S. W., … Bastien, C. H. (2009). The relationship of lead neurotoxicity to the eventrelated potential P3B component in Inuit children von arctic Québec. *Neurotoxicology, 30*, 1070–1077. https://doi.org/10.1016/j.neuro.2009.06.008

Chiplonkar, S. A., & Kawade, R. (2014). Linkages of biomarkers of zinc with cognitive performance and taste acuity in adolescent girls. *International Journal of Food Science and Nutrition, 65*, 399–415. https://doi.org/10.3109/09637486.2014.880667

Colom, R., Jung, R., & Haier, R. J. (2006). Distributed brain sites for the g-factor of intelligence. *Neuroimage, 31*(3), 1359–1365. https://doi.org/10.1016/j.neuroimage.2006.01.006

Costenbader, V., & Ngari, S. M. (2001). A Kenya standardisation of the Coloured progressive matrices. *School Psychology International, 22*, 258–268.

Counter, S. A., Buchanan, L. H., Rosas, H. D., & Ortega, F. (1998). Neurocognitive effects of chronic lead intoxication in Andean children. *Journal of Neurological Sciences, 160*, 47–53.

Dendir, S. (2013). Children's endowment, schooling, and work in Ethiopia. *WIDER working paper, no. 2013/086*. Information Center for Economics.

Dobrean, A., Raven, J., Comşa, M., Rusu, C., & Balazsi, R. (2008). Romanian standardization of Raven's standard progressive matrices plus. In J. Raven, & J. C. Raven (Eds.), *2008. Uses and Abuses of Intelligence*. Royal Fireworks Press.

Dutton, E. (2018). *J. Philippe Rushton: A life history perspective*. Thomas Edward Press.

Dutton, E., Bakhiet, S. F., Essa, Y. A. S., Blahmar, T. A., & Hakami, S. M. A. (2017). A negative Flynn effect in Kuwait: The same effect as in Europe but with seemingly different causes. *Personality and Individual Differences, 114*, 69–72. https://doi.org/10.1016/j.paid.2017.03.060

Dutton, E., Becker, D., Osman, H. H., 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. https://doi.org/10.1016/j.paid.2018.02.018

Figueredo, A. J., Cabeza de Baca, T., & Woodley, M. A. (2013). The measurement of human life history strategy. *Personality and Individual Differences, 55*, 251–255. https://doi.org/10.1016/j.paid.2012.04.033

Flynn, J. R. (2012). *Are we getting smarter? Rising IQ in the twenty-first century*. Cambridge University Press.

Fox, M. C., & Mitchum, A. L. (2013). A knowledge based theory of rising scores on "culture-free" tests. *Journal of Experimental Psychology: General, 142*, 979–1000. https://doi.org/10.1037/a0030155

Galler, J. R., Ramsey, F. C., Forde, V., & Archer, E. (1987). Long-term effects of early kwashiorkor compared with marasmus. II. Intellectual performance. *Journal of Pediatric Astroenterology & Nutrition, 6*, 847–854.

Galván, M., Uauy, R., López-Rodríguez, G., & Kain, J. (2013). Association between childhood obesity, cognitive development, physical fitness and social-emotional wellbeing in a transitional economy. *Annals of Human Biology, 4*, 101–106. https://doi.org/10.3109/03014460.2013.841288

Glewwe, P., & Jaccoby, H. (1992). *Estimating the determinants of cognitive achievement in low income countries*. Washington, DC: World Bank.

Grantham-McGregor, S., Powell, C., Walker, S., Chang, S., & Fletcher, P. (1994). The long-term follow-up of severely malnourished children who participated in an intervention program. *Child Development, 65*, 428–439.

Grantham-McGregor, S. M., Walker, S. P., Chang, S. M., & Powell, C. A. (1997). Effects of early childhood supplementation with and without stimulation on later development in stunted Jamaican children. *American Journal of Clinical Nutrition, 66*, 247–253.

Grigoriev, A., & Lynn, R. (2014). A study of the IQ in Kazakhstan. *Intelligence, 46*, 40–46. https://doi.org/10.1016/j.intell.2014.05.004

Haier, R. J., Colom, R., Schroeder, D. H., Condon, C. A., Tang, C., Eaves, E., & Head, K. (2009). Gray matter and intelligence factors: Is there a neuro-g? *Intelligence, 37*(2), 136–144. https://doi.org/10.1016/j.intell.2008.10.011

Hamadani, J. D., Tofail, F., Nermell, B., Gardner, R., Shiraji, S., Bottai, M., … Vahter, M. (2011). Critical windows of exposure for arsenic-associated impairment of cognitive function in pre-school girls and boys:A population-based cohort study. *International Journal of Epidemioliology, 40*, 1593–1604. https://doi.org/10.1093/ije/dyr176

Harari, R., Julvez, J., Murata, K., Barr, D., Bellinger, D. C., Debes, F., & Grandjean, P. (2010). Neurobehavioral deficits and increased blood pressure in school-age children prenatally exposed to pesticides. *Environmental Health Perspectives, 118*, 890–896. https://doi.org/10.1289/ehp.0901582

Heady, C. (2003). The effect of child labor on learning achievement. *World Development, 31*, 385–398. https://doi.org/10.1016/S0305-750X(02)00186-9

Heilmann, S. (2013). *Life-chances of children in Indonesia: The links between parental resources and children's outcomes in the areas of nutrition, cognition and health. A thesis submitted to the Department of Social Policy of the London School of Economics for the degree of doctor of philosophy, London. September 2013*.

Humble, S., Dixon, P., & Schagen, I. (2016). Assessing intellectual potential in Tanzanian children in poor areas of Dar Es Salaam. *Assessment in Education: Principles, Policy & Practice.*. https://doi.org/10.1080/0969594X.2016.1194257

Hunt, E., & Sternberg, R. J. (2006). Sorry, wrong numbers: An analysis of a study of a correlation between skin color and IQ. *Intelligence, 34*, 131–137. https://doi.org/10.1016/j.in-tell.2005.04.004

Hur, Y.-M., & Lynn, R. (2013). Twin-singleton differences in cognitive abilities in a sample of Africans in Nigeria. *Twin Research and Human Genetics, 16*, 1–8. https://doi.org/10.1017/thg.2013.44

Husain, N. I. A. E., Meisenberg, G., Becker, D., Bakhiet, S. F., Essa, Y. A. S., Lynn, R., & Al Julayghim, F. M. H. (2019). Intelligence, family income and parental education in the Sudan. *Intelligence, 77*, Article 101402. https://doi.org/10.1016/j.intell.2019.101402

Ijarotimi, O. S., & Ijadunola, K. T. (2007). Nutritional status and intelligence quotient of primary school children in Akure community of Ondo state, Nigeria. *Tanzania Health Research Bulletin, 9*, 69–76. https://doi.org/10.4314/thrb.v9i2.14307

Iloh, K. K., Ubesie, A. C., & Iloh, O. N. (2017). Do socio-demographic characteristics influence the performance of children on Raven progressive matrices in Enugu, Nigeria? *International Neuropsychiatric Disease Journal, 9*, 1–6. https://doi.org/10.9734//INDJ/2017/32267

Jardim-Botelho, A., Raff, S., Rodrigues, R. A., Hoffman, H. J., Diemert, D. J., Corrêa-Oliveira, R., … Gazzinelli, M. F. (2008). Hookworm, ascaris lumbricoides infection and polyparasitism associated with poor cognitive performance in Brazilian schoolchildren. *Tropical Medicine and International Health, 13*, 994–1004. https://doi.org/10.1111/j.1365-3156.2008.02103.x

Jensen, A. R. (1972). *Genetics and education*. Methuen.

Jensen, A. R. (1974a). Cumulative deficit: A testable hypothesis? *Developmental Psychology, 10*(6), 996–1019. https://doi.org/10.1037/h0037246

Jensen, A. R. (1974b). How biased are culture-loaded tests? *Genetic Psychology Monographs, 90*(2), 185–244.

Jensen, A. R. (1977). Cumulative deficit in IQ of blacks in the rural south. *Developmental Psychology, 13*(3), 184–191. https://doi.org/10.1037/0012-1649.13.3.184

Jensen, A. R. (1980). *Bias in mental testing*. Methuen.

Jensen, A. R. (1993). Psychometric g and achievement. In B. R. Gifford (Ed.), *Policy perspectives on educational testing* (pp. 117–227). Kluwer Academic Publishers.

Jukes, M. C. H., & Grigorenko, E. L. (2010). Assessment of cognitive abilities in multiethnic countries: The case of the Wolof and Mandinka in Gambia. *British Journal of Educational Psychology, 80*, 77–98. https://doi.org/10.1348/000709909X475055

Jukes, M. C. H., Pinder, M., Grigorenko, E. L., Smith, H. B., Walraven, G., & Bariau, E. M. (2006). Long-term impact of malaria chemoprophylaxis on cognitive abilities and educational attainment: Follow-up of a controlled trial. *PLoS Clinical Trials, 19*, 1–8. https://doi.org/10.1371/journal.pctr.0010019

Kitsao-Wekulo, P. K., Holding, P. A., Taylor, H. G., Abubakar, A., & Connolly, K. (2013). Neuropsychological testing in a rural African school-age population: Evaluating contributions to variability in test performance. *Assessment, 20*, 776–784. https://doi.org/10.1177/1073191112457408

Knowles, J. (2008). *Assessment of non-verbal intelligence in south African schools: Do language and gender bias performance on the Raven's standard progressive matrices? Master thesis at the university of the Witwatersrand, Johannesburg*.

Leon, C. A., Jaramillo, R., Martinez, S., Fernandez, F., Tollez, H., Lasso, B., & de Guzmdn, R. (1975). Sequelae of Venezuelan equine encephalitis in humans: A four year follow-up. *International Journal of Epidemiology, 4*, 131–140.

Li, S.-C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive pro-cesses across the life span. *Psychological Science, 15*(3), 155–163. https://doi.org/10.1111/j.0956-7976.2004.01503003.x

Lim, S. S., Updike, R. L., Kaldjian, A. S., Barber, R. M., Cowling, K., York, H., … Murray, C. J. L. (2018). Measuring human capital: A systematic analysis of 195 countries and territories, 1990–2016. *Lancet, 392*(10154), 1217–1234. https://doi.org/10.1016/S0140-6736(18)31941-X

Lohman, D. F., & Hagen, E. P. (2002). *Cognitive abilities test - form 6, Level H*. Riverside Publishing.

Lüdtke, O., Robitzsch, A., Trautwein, U., & Köller, U. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung - Probleme und Lösungen. *Psychologische Rundschau, 58*(2), 103–117. https://doi.org/10.1026/0033-3042.58.2.103

Lynn, R., & Becker, D. (2019). *The intelligence of nations*. Ulster Institute for Social Research.

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

Lynn, R., & Vanhanen, T. (2006). *IQ and global inequality*. Washington Summit Publishers.

Lynn, R., & Vanhanen, T. (2012). *Intelligence: A unifying construct for the social sciences*. Ulster Institute for Social Research.

Malda, M., van der Vijver, F. J. R., & Temaine, Q. M. (2010). Rugby versus soccer in South Africa: Content familiarity contributes to cross-cultural differences in test scores. *Intelligence, 38*, 582–595. https://doi.org/10.1016/j.intell.2010.07.004

Meisenberg, G. (2014). Cognitive human capital and economic growth in the 21st century. In T. Abrahams (Ed.), *Economic growth in the 21st century: Perspectives, role of governmental policies, potential and constraints* (pp. 49–106). Nova Publishers.

Menezes-Filho, J. A., Novaes, C., Moreira, J. C., Sarcinelli, P. N., & Mergler, D. (2011). Elevated manganese and cognitive performance in school-aged children and their mothers. *Environmental Research, 111*, 156–163. https://doi.org/10.1016/j.envres.2010.09.006

de Neubourg, E., & de Neubourg, C. (2011). *The impact of malnutrition and post traumatic stress disorder on the performance of working memory in children*. Preliminary Draft.

Nga, T. T., Winichagoon, P., Dijkhuizen, M. A., Khan, N. C., Wasantwisut, E., & Wieringa, F. T. (2011). Decreased parasite load and improved cognitive outcomes caused by deworming and consumption of multi-micronutrient fortified biscuits in rural Vietnamese schoolchildren. *The American Journal of Tropical Medicine and Hygiene, 85*, 333–340. https://doi.org/10.4269/ajtmh.2011.10-0651

Nokes, C., Grantham-McGregor, S. M., Sawyer, A. W., Cooper, E. S., & Bundy, D. A. P. (1992). Parasitic helminth infection and cognitive function in school children. *Proceedings of the Royal Society, Biological Sciences, 247*, 77–81.

Nokes, C., Grantham-McGregor, S. M., Sawyer, A. W., Cooper, E. S., Robinson, B. A., & Bundy, D. A. (1992). Moderate to heavy infections of Trichuris trichiura affect cognitive function in Jamaican school children. *Parasitology, 104*, 539–547.

Osman, H. A., Bakhiet, S. F. A., Alnaier, N. A. G., & Lynn, R. (2017). A study of the intelligence of south Sudanese refugee children. *Mankind Quarterly, 57*, 585–589.

Owen, K. (1992). The suitability of Raven's progressive matrices for various groups in South Africa. *Personality and Individual Differences, 13*(2), 149–159. https://doi.org/10.1016/0191-8869(92)90037-P

Peixoto, B., & Kalei, I. (2013). Neurocognitive sequelae of cerebral malaria in adults: A pilot study in Benguela central hospital, Angola. *Asian Pacific Journal of Tropical Biomedicine, 3*, 532–535. https://doi.org/10.1016/S2221-1691(13)60108-2

Petermann, F., & Petermann, U. (2014). *Wechsler intelligence scale for children - fourth edition, manual 1 - Grundlagen, Testauswertung und interpretation, Übersetzung und Adaptation der WISC-IV von David Wechsler*. Pearson Assessment.

Posner, M., & Barbey, A. K. (2020). General intelligence in the age of neuroimaging. *Trends in Neuroscience and Education, 18*, Article 100126. https://doi.org/10.1016/j.tine.2020.100126

Raven, J. (1981). *Irish and British Standardisations*. Oxford Psychologist Press.

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology, 41*, 1–48.

Raven, J. (2008a). *Coloured progressive matrices and Crichton vocabulary scale*. Pearson.

Raven, J. (2008b). *Standard progressive matrices and Mill Hill vocabulary scale*. Pearson.

Raven, J., Raven, J. C., & Court, J. H. (1985). *Manual for Raven's progressive matrices and vocabulary scales*. H.K: Lewis & Co.Ltd.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford Psychologist Press.

Raven, J., Raven, J. C., & Court, J. H. (1999). *Manual zu Raven's progressive matrices und vocabulary scales*. Swets & Zeitlinger B. V.

Rindermann, H., Becker, D., & Coyle, T. R. (2016). Survey of expert opinion on intelligence: Causes of international differences in cognitive ability tests. *Frontiers in Psychology, 7*(399), 1–9. https://doi.org/10.3389/fpsyg.2016.00399

Rodríguez, T. (2012). *Environmental pesticide exposure and neurobehavioral effects among children of Nicaraguan agricultural workers. Ph.D thesis, University of Uppsala. Acta Universitatis Upsaliensis. 820*. Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine.

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

Soemantri, A. G. (1989). Preliminary findings on iron supplementation and learning achievement of rural Indonesian children. *American Journal of Clinical Nutrition, 50*, 698–702.

Spearman, C. E. (1946). Theory of the general factor. *British Journal of Psychology, 36*, 117–131.

Stacey, C. L., & Carleton, F. O. (1955). The relationship between Raven's Coloured progressive matrices and two tests of general intelligence. *Journal of Clinical Psychology, 11*, 84–85.

Stein, A. D., Behrman, J. R., DiGirolamo, A., Grajeda, R., Martorell, R., Quisumbing, A., & Ramakrishnan, U. (2005). Schooling, educational achievement, and cognitive functioning among young Guatemalan adults. *Food and Nutrition Bulletin, 26*, 46–54. https://doi.org/10.1177/15648265050262S105

Tadmouri, G. O., Nair, P., Obeid, T., Al Ali, M. T., Al Khaja, N., & Hamamy, H. A. (2009). Consanguinity and reproductive health among Arabs. *Reproductive Health, 6*, 17. https://doi.org/10.1186/1742-4755-6-17

United Nations Development Programme. (2017). Human Development Reports. http://hdr.undp.org/en/composite/HDI4.

Van de Vijver, F. J. R., & Brouwers, S. A. (2009). Schooling and basic aspects of intelligence: A natural quasi-experiment in Malawi. *Journal of Applied Developmental Psychology, 30*(2), 67–74. https://doi.org/10.1016/j.appdev.2008.10.010

Vernon, P. E., & Parry, J. B. (1949). *Personnel selection in the British forces*. University of London Press.

Volken, T. (2003). IQ and the wealth of nations: A critique of Richard Lynn and Tatu Vanhanen's book. *European Sociological Review, 19*(4), 411–412. https://doi.org/10.1093/esr/19.4.411

Warne, R. (2023). National mean IQ estimates: Validity, data quality, and recommendations. *Evolutionary Psychological Science, 9*, 197–223. https://doi.org/10.1007/s40806-022-00351-y

Warne, R. T., & Burningham, C. (2019). Spearman's g found in 31 non-western nations: Strong evidence that g is a universal phenomenon. *Psychological Bulletin, 145*(3), 237–272. https://doi.org/10.1037/bul0000184

Wasserman, G. A., Liu, X., Parvez, F., Ahsan, F., Factor-Litvak, P., Kline, J., … Graziano, J. H. (2007). Water arsenic exposure and intellectual function in 6-year-old children in Araihazar, Bangladesh. *Environmnetal Health Perspectives, 115*(2), 285–296. https://doi.org/10.1289/ehp.9501

Weihe, P., Hansen, J. C., Murata, K., Debes, F., Jørgensen, P. J., Steuerwald, U., … Grandjean, P. (2002). Neurobehavioral performance of Inuit children with increased prenatal exposure to methylmercury. *International Journal of Circumpolar Health, 61*, 41–49.

Weiß, R., & Osterland, J. (1980). *Grundintelligenztest Skala 1 - CFT 1. Handanweisung für Durchführung, Auswertung und Interpretation*. Westermann.

Weiß, R., & Osterland, J. (1997). *Grundintelligenztest Skala 1 - CFT 1. Handanweisung für Durchführung, Auswertung und Interpretation*. Hogrefe.

Weiß, R., & Osterland, J. (2013). *CFT 1-R Grundintelligenztest, Skala 1 revision*. Hogrefe.

Woodley, M. A. (2009). Inbreeding depression and IQ in a study of 72 countries. *Intelligence, 37*, 268–276. https://doi.org/10.1016/j.intell.2008.10.007

Žebec, M., Demetriou, A., & Kotrla-Topić, M. (2015). Changing expressions of general intelligence in development: A 2-wave longitudinal study from 7 to 18 years of age. *Intelligence, 49*, 94–109. https://doi.org/10.1016/j.intell.2015.01.004