Text: 4818 words
Tables: 3
Figure: 3

Seventy years, 1,000 samples, and 300,000 SPM scores: A new meta-analysis of Flynn effect patterns

Peera Wongupparaj, [a,b,d*] Rangsirat Wongupparaj[c], Robin G Morris[d], Veena Kumari[e]

[a]Cognitive Science and Innovation Research Unit (CSIRU), College of Research Methodology and Cognitive Science, Burapha University, Thailand

[b]Thai Clinical Psychologist Association, Thailand

[c]Department of Curriculum and Instruction, Faculty of Education, Chulalongkorn University, Thailand

[d]Department of Psychology, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, UK

[e]Centre for Cognitive Neuroscience (CCN), College of Health, Medicine and Life Sciences, Brunel University London, UK

*Correspondence to:

Peera Wongupparaj

Cognitive Science and Innovation Research Unit (CSIRU),

College of Research Methodology and Cognitive Science, Burapha University, Thailand 20131

Tel: 061 120 7302, Email: Peera.1.wongupparaj@kcl.ac.uk; peera.wo@go.buu.ac.th

Abstract

Several studies have investigated and found the gradual rise in IQ over time or the Flynn effect (FE) but inconsistent results on the FEs across types of countries and age groups were reported. The current cross-temporal meta-analysis aimed to examine the temporal correlations between mean IQ scores from Standard Progressive Matrices and year of publication, moderated by types of countries and age groups covering seven decades (1948-2020). The given relationships were weighted by sample sizes. The dataset included 1038 independent samples (N=299155) from 72 countries. The results generally supported the FE with the IQ gain of 0.22 points per year, but the magnitudes of the IQ gains depended on types of countries and age groups. Stronger FEs were evident in middle-income countries and younger generations. A multicausal explanatory framework should be utilized to explain the underlying mechanism of the secular IQ gains across factors.

Keywords: Cross-Temporal Meta-Analysis; the sample-size weighted-regression equation; Standard Progressive Matrices; Non-verbal intelligence

1. Introduction

The well-known Flynn effect (FE) refers to an observed rise in average IQ scores that occurs over time, in the order of decades, with each emerging generation showing increased performance on IQ tests (Pietschnig & Voracek, 2015; Rodgers, 1998; te Nijenhuis & van der Flier, 2013; Trahan et al., 2014). This effect was originally found in North American populations, where James Flynn conducted large-scale studies that showed substantial gains in mean IQ scores, with an increase of three IQ points per decade in full-scale IQ scores (Flynn, 1984, 1987). Since then, there has been a plethora of studies investigating the FE, with over 371 studies documented in 2023 (Scopus, 2023). The presence of the FE has led to educational, psychological and social controversies, an example being the extent to which rising IQ scores change the threshold for learning disability diagnosis such that a person with intellectual or developmental disabilities might no longer qualify for social security disability benefits (Kanaya, 2016; Reardon, 2014). Likewise, in countries where the death penalty still exists, such as the USA, IQ scores can be used to help threshold those who may be thought inappropriate for the death penalty, specifically people who meet the criteria for learning disabilities (Flynn & Widaman, 2008). This is one reason why IQ test manufacturers revise and publish new norms every 10 to 15 years, to avoid obsolete norms and preserve the generational validity of IQ testing (Kanaya, 2016).

Since the Flynn effect was first observed, there has been speculation about causation. Dickens and Flynn (2001) proposed that complex gene-environment interactions influence IQ gains and recent research has supported the FE being moderated by the interplay of genetic and environmental drivers (Giangrande et al., 2022). Nonetheless, previous studies have suggested that various direct potential environmental mechanisms appear to be more responsible for changing average IQ scores (Beller et al., 2022; Bratsberg & Rogeberg, 2018). These include the following factors: Better nutrition and improved standards of living (Lynn, 2009; Martorell, 1998); improved public health systems, including vaccines and less exposure to harmful pathogens and disease (Clark et al., 2016; Eppig et al., 2010; Galiciolli et al., 2022; Kaufman et al., 2014); improved education and school systems, including better education within families and having more highly educated parents (Baker et al., 2015; Brinch & Galloway, 2012); better home environments in terms of cognitive stimulation and emotional support (O'Keefe & Rodgers, 2022); and more-intelligent social environments and increased environmental complexity including computers, smart phones, TV and internet (Bordone et al., 2015; Bratsberg & Rogeberg, 2018; Clark et al., 2016; Schooler, 1998). Accordingly, such findings compellingly support a multicausal explanatory framework for the FE that includes environmental or social multipliers (Pietschnig, 2016).

Whilst investigation of environmental factors may provide explanations of the FE, other investigations have focused on the demographic characteristics of populations to determine their impact on the size of the FE. In this study, we have used a very large sample meta-analytical approach to investigate further the effects of two main demographic variables, namely the degree of development or wealth of the country of origin and the age of the specific cohorts used in studies. These are considered here in turn:

(1) Country Development or Wealth: The FE is generally documented and replicated across countries using cohort and meta-analysis approaches (Brouwers et al., 2009; Pietschnig et al., 2010). The FE and economic growth in 28 countries have been investigated and significant positive correlation coefficients between cognitive ability and wealth have been observed (Rindermann & Becker, 2018). Specifically, a one-point increase in a nation's mean IQ was related to a continuing 0.11% annual increase in Gross Domestic Product (GDP) per capita (Jones & Schneider, 2006).

Likewise, a larger FE was observed in emerging European countries and positive associations between the FE and GDP was found (Weber et al., 2017). The growing economies in developing countries and rising mean IQ scores over a period of six decades were also evident (Wongupparaj et al., 2015; Wongupparaj et al., 2017). Furthermore, the estimated mean IQ scores for developed nations outperformed those of developing nations in many studies (Brouwers et al., 2009; Wicherts et al., 2010; Wongupparaj et al., 2015).

(2) Age Effects: Generational IQ changes over time among individuals in the same age group has been reported and, as suggested above, ascribed mainly to the environmental factors (e.g., see Pietschnig & Voracek, 2015 for a review). Several studies have investigated the FE across the lifespan finding the effect robust (Rönnlund & Nilsson, 2009; Shakeel & Peterson, 2022; Wongupparaj et al., 2015). Nevertheless, studies concerning the relative strength of the FE at different stages of development have been conflicting.One meta-analysis suggested a larger FE in adults than in children (Pietschnig & Voracek, 2015). However, another large cross-sectional study observed a larger FE in more recent generations (Merten et al., 2022), and yet another meta-analysis showed no differences across age groups (Trahan et al., 2014).

Given the universality of this effect and consequent global implications, Rodgers (2015) has suggested the field seems poised for confirmatory analysis. A large-scale meta-analysis covering several decades, a comparable number and type of country, and massive sample sizes may further establish the extent of IQ gains due to environmental potency between generations.

In the light of the evidentiary and theoretical bases concerning a focal determinant of the global FE, a clearer understanding of secular gains has several potential implications for both exploratory and confirmatory efforts. Here we present what we believe is the largest study of the FE on fluid intelligence to date, comprising 1,038 independent samples and 299,155 total participants from 71 countries to examine the FEs over a period of seven decades (1948-2020). Type of country (i.e., low-, middle-, and high-income) and age group (i.e., <12 years old, 13-19 years old, 20-39 years old and >40 years old) were investigated as moderators of the secular rise of average IQ scores.

2. Method

2.1 Literature search

Prior meta-analyses studies guided the methodology, the literature search and search terms, inclusion and exclusion criteria, and coding (Pietschnig et al., 2010; Wongupparaj et al., 2015; Wongupparaj et al., 2017). The target articles were systematically located through 11 scientific databases, namely, i) ScienceDirect; ii) SpringerLink; iii) PsycArticles; iv) PsycInfo; v) BMJ; vi) Frontiers; vii) PlosOne; viii) BioMed Central; ix) MDPI journals; x) ProQuest dissertations and theses, and xi) Open Access Theses and Dissertations (OATD). The study focused on investigations using the Standard Progressive Matrices (SPM), this being the most universally used IQ test across the globe over time. It was developed in 1936 and considered to be a culturally fair test and a well-validated measure of general ability or fluid intelligence (Feis, 2010). The mean scores of the SPM and relevant parameters (i.e., standard deviation from mean, reliability, and test version) were collected and analyzed. The SPM has been used extensively in research contexts from early childhood to old age (8- to 65-year-olds) both in English and non-English speakers since 1962, and thus it is useful for comparative studies (Raven, 2000). The SPM versions adopted similar test instructions, materials, scoring procedures, and principles so that it ensured null test version effects on the FE. The following search terms were employed individually and in combination using the Boolean OR function to maximize search sensitivity; "Standard Progressive Matrices" and "Raven's Progressive Matrices." To retrieve the grey literature or unpublished findings, thesis/dissertation and in-press articles were also searched.

2.2 Inclusion and exclusion criteria

All English literature searches were screened and coded by two researchers. Raters resolved any coding disagreements via discussion and consensus. Independent studies were included in the meta-analysis if these studies provided the mean and/or standard deviation raw scores of the SPM sets A to E. Studies were excluded following an eligibility check on full-text articles if they reported the mean and/or standard deviation only of shortened (e.g., nine-item version), odd- or even-item, combined (e.g., SPM 38 or the Combined Raven's Test for Children: CRT-C) or modified versions (e.g., the first two items of set A were employed for practice). The standardized, mean, median and geometric mean scores were excluded. The mean and/or standard deviation raw scores of other SPM variants (i.e., SPM Plus and 2 versions) were also excluded. In addition, any two-choice answers for each SPM item, instead of the four to six choices in the original version, were removed.

Moreover, if studies adopted a test–retest approach, only the SPM mean and/or standard deviation scores for pre-test or baseline were collected and if various articles examined the same sample or used the same dataset, the SPM means and standard deviations were treated as a single data point. Studies were removed if they investigated clinical research participants (e.g. with mental or physical disabilities or difficulties) except when they also reported data for control groups or healthy controls, as such data are appropriate for use in the current study. Furthermore, review articles, systematic reviews, meta-analyses, research protocols, letters, personal communications, and case reports were excluded because no targeted statistical parameters are reported in these studies. Finally, the year of data collection was coded as two years before year of publication, unless the data collection date was specifically mentioned in the literature, in line with previous cross-temporal meta-analyses (Oliver & Hyde, 1993; Twenge et al., 2010). Accordingly, the target parameters of this study were collected between 1948 and 2020.

2.3 Moderator variables

To investigate possible moderators influencing the mean SPM raw scores other than year of data collection, two researchers independently coded the following variables: Type of country and age group. Type of country was divided into “low-” “middle-” or “high-” income in line with the World Bank Atlas (2022) (Lists of 14 low-, 25 middle-, and 32 high-income countries are included in the final dataset; see **Figure 1** and **Appendix A).** Lastly, the mean age from each independent study was collected and inputted into the final dataset. To further investigate the specific FE for each age group, mean ages were ordered and divided into four age groups (i.e., <12 years old, 13-19 years old, 20-39 years old and >40 years old) (Wongupparaj et al., 2015). All targeted variables and characteristics are shown in **Table 1.**

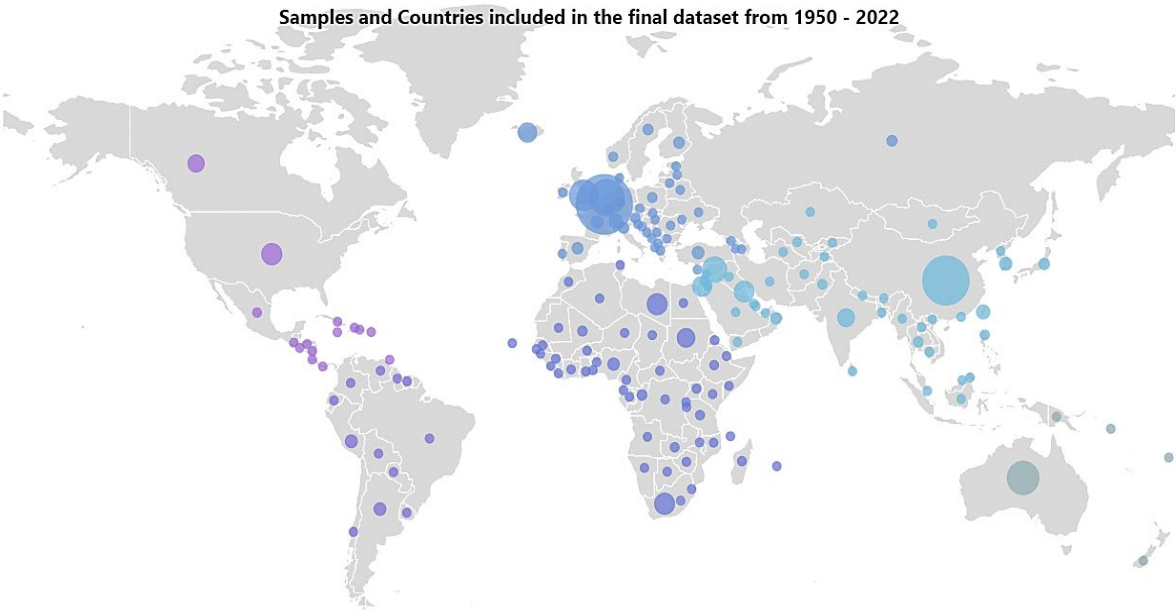


Figure 1. Sample sizes across 71 countries (14 low-, 25 middle-, and 32 high-income countries) included in the final dataset. The bubble sizes provide a visual analogue concerning the relative sample sizes and the colors represent different continents for the samples.

2.4 Final sample

The search strategies and study selection were conducted in line with the PRISMA guidelines and the PRISMA flow chart is demonstrated in **Figure 2.** The 11-electronic database search produced 10,594 studies, incorporating 10,273 studies from published sources and 321 studies from unpublished sources or grey literature. Of these, unrelated and duplicated studies were removed after initial screening. The EndNote version X10 was used to identify and remove duplicated titles. Accordingly, 7,203 studies remained for the next step; out of these, 2,953 studies were excluded after title, abstract, and keyword reviews. Subsequently, only 413 independent studies (N) with 1,038 independent samples (k) fulfilled the inclusion and exclusion criteria, resulting in 299,155 participants (n) in total.

Table 1 Sample characteristics.

Variables	Mean	SD	Skewness	Kurtosis	Min-Max	Number of independent samples (k)	Participant Totals (n)
Year of data collection	2003.74	13.91	-1.48	2.17	1948-2020	1,038	299,155
Mean age	18.37	13.04	2.61	7.53	4-79.10	1,010	296,390
Mean SPM score	38.77	10.46	-0.48	-0.37	9.50-57.85	1,038	299,155
SD SPM score	7.32	2.90	0.75	4.14	0.45-26.50	869	270,772
Mean IQ score	100	15	-0.48	-0.37	58.03-127.39	1,038	299,155
Sample size	287.80	1122.27	15.11[a]	264.73[a]	6-21,432	1,037	298,452
Type of country							
Low-income						118 (11.4%)	49,498
Middle-income						271 (26.1%)	61,173
High-income						649 (62.5%)	187,781
Age group							
<12 years old						411 (39.6%)	110,468
13-19 years old						297 (28.6%)	100,685
20-39 years old						239 (23%)	33,514
>40 years old						63 (6.1%)	11,132

Note: [a]Log-transformation was employed (sample size$_{\text{skewness and kurtosis}}$ = 0.50 and -0.29).

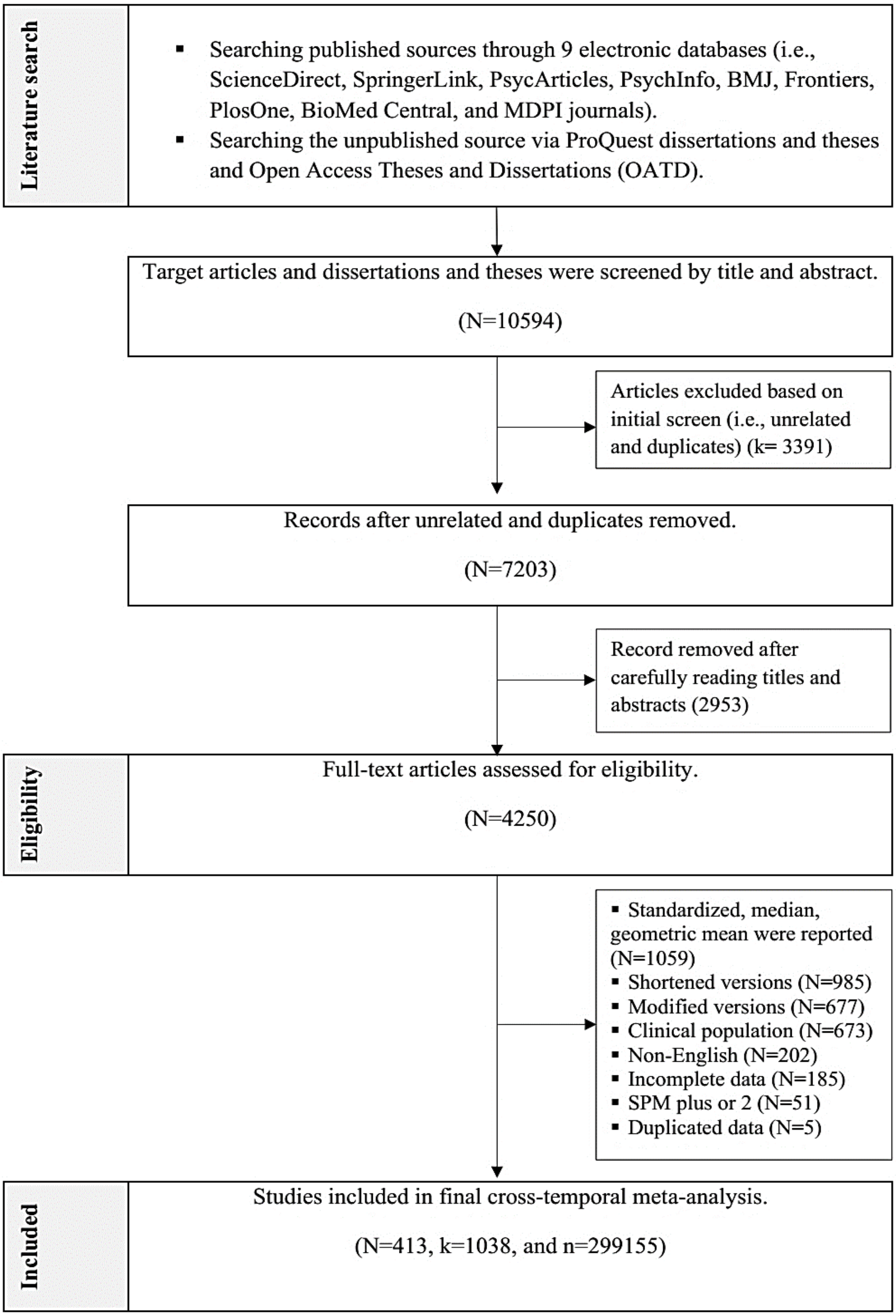


Figure 2. PRISMA flow chart of the search strategies and study selection. _Note:_ N denotes independent studies, k represents independent samples, and n denotes sample sizes.

2.5 Statistical analyses

The current study used Cross-Temporal Meta-Analysis (CTMA) to examine the extent to which the IQ scores changed over decades through using temporal correlation coefficients, and utilizing year of data collection as a predictor (Twenge, 2011; Twenge et al., 2010). Bivariate least squares linear regressions were calculated for each predictor, and weighted by the sample size for each study. The weighted least squares estimation with sample-size weights was employed such that the studies with larger sample sizes were assigned greater weights because they presumably were more representative and more closely approximated the mean of the population (Donnelly & Twenge, 2017; Lipsey & Wilson, 2001). The mean age of participants and type of countries were inputted into the regression model as moderators.

To determine an overall size of the FE, the magnitude of score changes was calculated using the sample-size weighted-regression equation with IQ scores as the outcome measure and the year of data collection as the predictor. Moreover, to compare the IQ score gain with previous meta-analytic studies, IQ scores were transformed from the raw SPM mean scores based on the linear transformations of z-scores from each independent study, using a = 15 and b =100, so that IQ = 15z + 100, with population mean (μ) IQ = 100, and population SD (σ) = 15. In addition, two different sample-size weighted-regression equations and IQ score gains were separately calculated between 1948-1985 and 1986-2020. More specifically, the specific FE gains for the two periods were estimated in line with the linear trends shown in **Figure 3** and also the main findings from Pietschnig & Voracek (2015) in that yearly fluid IQ gains rose by 0.43 points until 1985 and increased by 0.22 points until 2013.

Analyses were conducted using the IBM SPSS version 28 and Medcalc. To mitigate publication biases, the current dataset contained the mean SPM scores from published and unpublished sources. Furthermore, the target mean scores for SPM rarely relate to the central hypotheses or objectives considered by the journal articles, tending to be reported on an incidental basis as descriptive information, sample characteristics, or supplementary data (Booth et al., 2016; Pietschnig et al., 2010; Schmidt & Hunter, 2015). Several descriptive data and sample characteristics were collected from each study, namely, sample size, means, SDs, and reliability; this is not typically the type of data that is subject to tests of statistical significance and accordingly would not be a main cause of failure to publish because of non-significance (Trahan et al., 2014).

3. Results

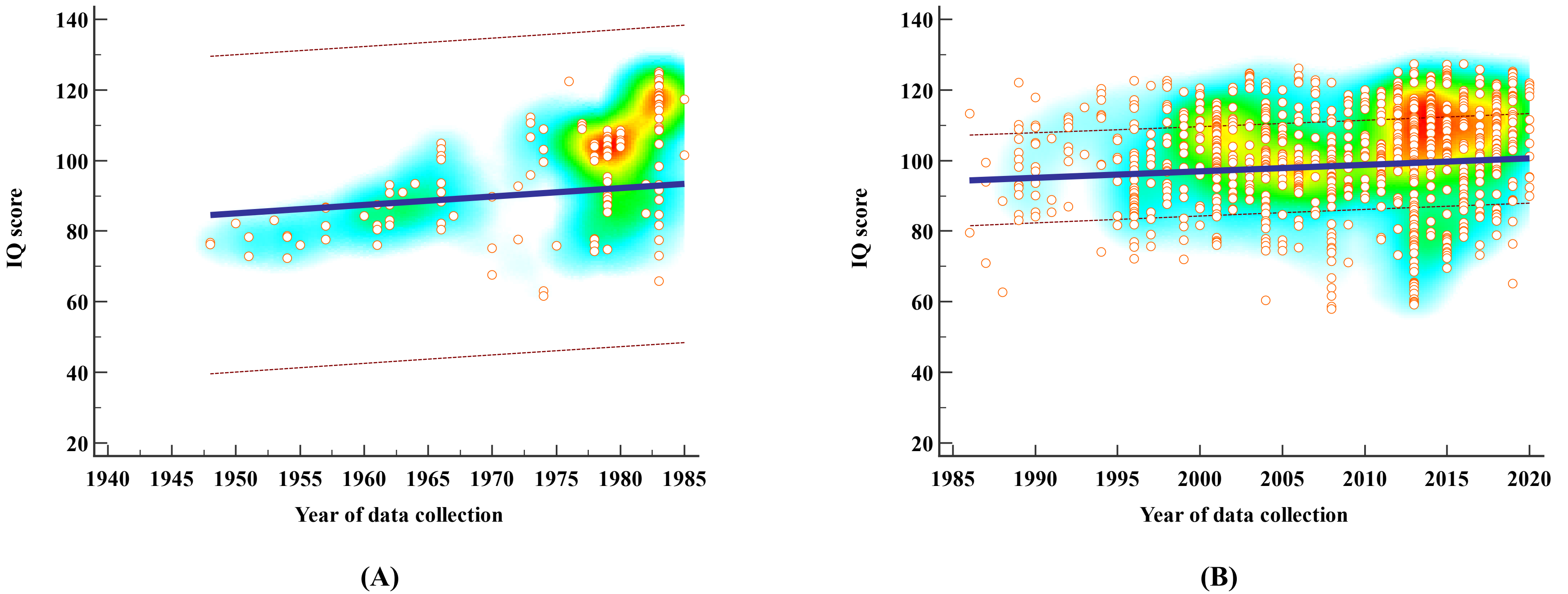
For the simple regression model, the SPM mean scores significantly increased between 1948 and 2020. In this model, the year of data collection could explain 14% of the variance in the IQ scores (F=173.74, df=1, p<.01) in terms of the R^2 value. Furthermore, the SPM mean scores showed a greater rise between 1948 and 1985 relative to that between 1986 and 2020. In addition, the year of data collection could explain around 6% and 1% of the variance in the IQ scores (F=8.36, df=1, p<.01 and F=11.52, df=1, p<.01) from 1948-1985 to 1986-2020 respectively (See **Table 2** and **Figure 3**).

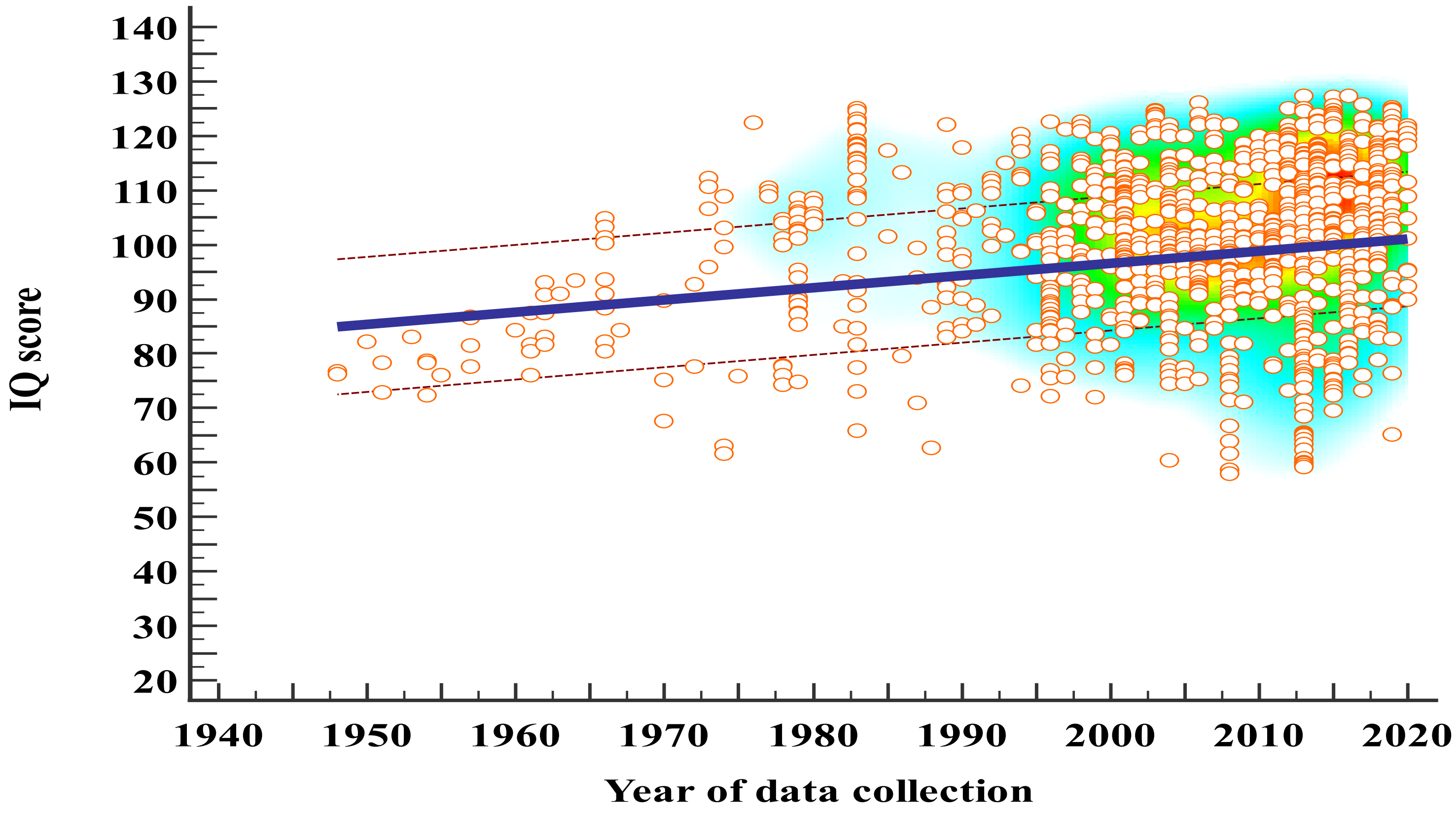
For the multiple regression model, the main effects of the year of data collection alongside type of country, and mean age significantly predicted the IQ scores. For interaction effects, type of country and mean age moderated the relationship between the year of data collection and the IQ scores. In addition, the year of data collection (β=.42, t=3.19, p<.01) exerted the strongest effect on the IQ scores, followed by the interaction effect between the year of data collection and mean age (*β*=-.34, *t*=-3.90, *p*<.01) and type of country (*β*=.28, *t*=8.94, *p*<.01), respectively. Moreover, the multiple predictors could account for 17% of the variance in the IQ scores (*F*=42.15, *df*=5, *p*<.01) in terms of the R^2 value. The results of these analyses are shown in **Table 2.**

Table 2. Single and multiple weighted least squares regression analyses for IQ scores.

Target variables	*b*	95% CI	SE	*β*	*t*	R^2
Single regression (1948-2020)						.14
Intercept	-351.56	-418.01 – -285.10	33.87	-	-10.38**	
Year of data collection	0.22	0.19 – 0.26	0.02	.38	13.18**	
Single regression (1948-1985)						.06
Intercept	-380.37	-701.71 – -59.03	162.35	-	-2.34*	
Year of data collection	0.24	0.08 – 0.40	0.08	.25	2.89**	
Single regression (1986-2020)						.01
Intercept	-270.38	-483.68 – -57.08	108.68	-	-2.49*	
Year of data collection	0.18	0.08 – 0.29	0.05	.11	3.39***	
Multiple regression (1948-2020)						.17
Intercept	-476.76	-821.12 – -132.40	175.48	-	-2.72**	
Year of data collection (YD)	0.28	0.11 – 0.45	0.09	.42	3.19**	
Mean age (MA)	0.15	0.09 – 0.22	0.03	.14	4.53**	
Type of country (TC)	5.01	3.91 – 6.11	0.56	.28	8.94**	
YD × MA	-0.01	-0.02 – -0.01	0.01	-.34	-3.90**	
YD × TC	0.05	0.01 – 0.11	0.03	.21	2.03*	

Note: Weighted by sample size; *<.05; **<.01.


IQ score
140
120
100
80
60
40
20
1940
1945
1950
1955
1960
1965
1970
1975
1980
1985
Year of data collection
(A)
IQ score
140
120
100
80
60
40
20
1985
1990
1995
2000
2005
2010
2015
2020
Year of data collection
(B)



(C)

Figure 3. Weighted simple regressions of the relationship between the IQ score and year of data collection during (A) 1948-1985, (B) 1986-2020, and (C) 1948-2020. The dash lines represent 95% confidence interval. The heat maps indicate density of independent samples.

Concerning the overall FEs across type of countries and age groups during 1948-2020, positive correlation coefficients were observed between the SPM mean scores and the year of publication (*r*=.38, *p*<.01) with an IQ gain score of 0.22 per year. Nonetheless, a more positive or stronger relationship between the SPM mean scores and the year of publication was found during 1948-1985 (*r*=.25, *p*<.01) than during 1986-2020 (*r*=.11, *p*<.01). Moreover, the specific FEs were 0.24 and 0.18 points per year from 1948-1985 to 1986-2020, respectively (see **Table 3).**

Specifically, stronger correlation coefficients were observed in middle-income countries (*r*=.60, *p*<.01) with the IQ gain score of 0.29 per year, relative to low- and high- income countries (*r*=.07, *ns* and *r*=.36, *p*<.01). The strongest and most positive direction of the correlation coefficient was observed in 13-19 years old (*r*=.45, *p*<.01) with an IQ gain score of 0.25 per year, followed by those aged 20-39 years old (*r*=.32, *p*<.01) with an IQ gain score of 0.23 per year, then those aged more than 40 years old (*r*=.26, *p*<.01) with an IQ gain score of 0.35 per year, and, finally, those aged less than 12 years old (*r*=.16, *p*<.01) with the IQ gain score of 0.12 per year, respectively (see **Table 3).**

Table 3. The correlation coefficients (*r*) between IQ score and year of data collection during 1948-2020 as weighted by sample size for type of country and age group.

Target variables (**Independent samples**)	***r***	**IQ gain per year (95% confidence interval)**
IQ score (1037) during 1948-2020	.38**	0.22 (0.19 – 0.26)
IQ score (126) during 1948-1985	.25**	0.24 (0.08 – 0.42)
IQ score (911) during 1986-2020	.11**	0.18 (0.07 – 0.29)
Type of countries during 1948-2020		
Low-income (118)	$.07^{ns}$	0.04 (0.05 – 0.13)
Middle-income (271)	.60**	0.29 (0.27 – 0.39)
High-income (648)	.36**	0.20 (0.16 – 0.24)
Age groups during 1948-2020		
<12 years old (411)	.16**	0.12 (0.05 – 0.18)
13-19 years old (297)	.45**	0.25 (0.20 – 0.31)
20-39 years old (238)	.32**	0.23 (0.14 – 0.32)
>40 years old (63)	.26*	0.35 (0.01 – 0.68)

Note: Weighted by sample size; *<.05; **<.01; ns: not significant.

4. Discussion

To the best of our knowledge, the present investigation provides the first analysis of fluid IQ changes as indexed by the SPM over seven decades, also being the largest known analysis of the FE, covering 71 countries, 1,038 independent samples and 299,155 as the total population. The IQ scores were derived from a standardized and most universally employed intelligence test. The SPM was originally developed in 1936 (Raven, 1936) and it has been regarded as one of the most widely used culturally reduced measures of non-verbal or fluid intelligence (Sternberg, 2019). Flynn (1987) also mentioned that in the case of the Raven's Progressive Matrices (RPM), the test items might not be easily learned from one generation to another. Further, the significance of IQ gains across generations depends on whether or not they manifest themselves on culturally reduced measures like the RPM since the RPM maximizes problem-solving abilities but minimizes the demand for more certain skills and knowledge of symbols and words (Flynn, 1987).

Due to the fact that the SPM has remained unaltered for more than seven decades, the sizable meta-analysis dataset from the current results generally support the IQ secular gain. Moreover, the present results yield many points of interest, including the FEs were observed across countries and age groups. For moderating effects on the IQ scores, the current findings suggested more robust IQ rises for middle-income countries and those in the teenage years (13-19 yrs).

The study highlights certain notable characteristics of the FE. First, it is obvious that the FE can be observed on the SPM, with our findings being generally clear and consistent with previously reported IQ gains using this measure (Colom et al., 1998; Wongupparaj et al., 2015). The IQ gain over 72 years (1948-2020) from the current study was 16.13 or 0.22 points per year with a lower and upper bound of 0.19 and 0.26 points. In addition, the stronger FE was evident during the first four decades (i.e., 1948-1985) and this observation was in line with Pietschnig & Voracek (2015). This current IQ gain over 72 years was lower than that found in the previous literature, with Flynn's study revealing a 0.30-point rise per year on the Stanford-Binet (SB) and Wechsler intelligence tests between 1932 and 1978 (Flynn, 1984). However, the gain rate per year of 0.22 points was similar to the analysis of Color Progressive Matrices (CPM) and SPM scores over 28 years (1980-2008) from ages between 5-15 years old with 0.22 points per year as recorded by Flynn (2012)

Using similar IQ measures, Pietschnig and Voracek (2015) also showed a fluid IQ gain of 0.41 points per year from 1924 to 2013. Further, this study also suggested that annual IQ gains increased by 0.43 points until 1985, then with weaker gains of 0.22 points until 2013. The slower IQ gain for the current results may be explained in terms of the generalizability of the sizable dataset that is beyond any specific type of country, age group, sample size, and research method (i.e., experimental, quasi-experimental, and non-experimental). Thus, the lower gain score may reflect the heterogenous populations covering seven decades of the SPM administration, with the FE effect likely changing in potency over time. Moreover, it is also evident that the current dataset included more than 912 independent samples (88%) from 1986 to 2020. Thus, the lower IQ gain may not conflict with the finding from Pietschnig and Voracek (2015), but reflect the greater sampling domain.

Second, by dividing the type of country into low-, middle-, and high-income countries, the current study provided empirical support for the previous studies suggesting that the SPM mean score for developed countries were greater than those of developing countries (Brouwers et al., 2009; Wicherts et al., 2010; Wongupparaj et al., 2015). It should be noted that the term 'developing country' is used here because of the historical labelling. Other distinctions such as in the present study dividing countries into low, middle and high income may more objectively and helpfully define country differences. Moreover, the present findings indicated nearly two-fold IQ gains for middle-income countries relative to high-income countries and a six-fold IQ rise in comparison to low-income countries. These results may reflect a negative FE in high-income countries over certain decades but also unimproved education and health systems in low-income countries (Orach, 2009; Seferidi et al., 2022). We intend to address these research questions in the future. A decline in FE or even anti-FE has been evident in seven high-income countries (Dutton et al., 2016; Vainikainen & Hautamäki, 2022). This study provides the first evidence suggesting that a steadily stronger

FE over decades for middle-income countries leads to higher IQ gain scores relative to those of high-income countries (0.29 vs 0.20).

The growing economies, improved labor skills, strengthened innovation capabilities as indexed by the GDP were positively linked to the IQ gains (Jones & Schneider, 2006; Rindermann & Becker, 2018; Weber et al., 2017). It should be noted that the income category of the countries is defined in this study relative to the current date and not historically. Hence, it is the case that categories will change in an upward trajectory, with the middle-income countries as defined here likely showing more critical change way from low income. The GDP growth of year 2021 supported this claim in that the high-income countries showed a GDP growth of 5.1 but the middle-income countries demonstrated a GDP growth of 6.9. Further, it may also highlight a dominant and complex gene-environment influences on IQ rises based on Giangrande et al. (2022) and the theory of individual and social multipliers (Dickens & Flynn, 2001). It was also hypothesized on the documented significant IQ gains in middle-income countries that it resonates best with increased parental education, smaller families, attitude towards schooling, and changes in children's nutrition and health (Bratsberg & Rogeberg, 2018; Woodley, 2011).

Third, the large size of the study enables it to further explore the generational IQ changes over decades of age groups using meta-analysis. The findings support the FE occurring across the lifespan affecting the age ranges from less than 12 yrs and to greater than 40 yrs. Here, the FE was most prominent in the 13-19 years old cohort, followed by the 20-39 years old cohort. The current evidence is consistent with an observed greater FE in more recent generations due to possible health-related or lifestyle changes and improved education (Merten et al., 2022; Pietschnig, 2016; Pietschnig & Voracek, 2015). Various potential environmental mechanisms have been proposed to explain the secular IQ rise across age groups, with agreement that a complex combination of several factors or a multicausal explanatory framework may alter by cohort and environmental context (Bratsberg & Rogeberg, 2018; Dickens & Flynn, 2001; Pietschnig, 2016). Increased exposure to technology has also been proposed as an underlying mechanism for the malleability of cognitive abilities (Pietschnig, 2016; Sauce et al., 2022).

In relation to the FE, the internet and gaming have been found to act as cognitive enhancers for non-verbal IQ in children (Kokkinakis et al., 2017; Sauce et al., 2022) and even enhance cognitive reserve in older adults born after 1941 (Berner et al., 2019; Ihle et al., 2020; Kim & Han, 2022; Li et al., 2022; Voinea et al., 2020). Internet usage has been rising across age groups in middle-income countries from 8.1% in 2005 to 57% in 2021 (Statista, 2022) as well as the growing gaming industries in middle-income countries (e.g., China) (Palma-Ruiz et al., 2022). The informal learning environment of internet and online gaming may support learners to develop a new profile of cognitive skills over time. The relatively large rise in non-verbal or visual intelligence may stem from the ever-more complex visual environments that involve the development of more sophisticated visuo-spatial skills (Greenfield, 2009).

Finally, the current meta-analysis aimed to produce less biased estimators for the FE across countries and age groups. With a pool of 299,155 participants, a variety of sample characteristics was incorporated into the dataset, including research methodology (i.e., non-experimental, quasi-experimental, and experimental designs), occupations (i.e., students, undergraduates, graduates, police officers, public servants, young offenders, conscripts, and farmers), years of education (from 1.50 to 16.65 years), and sample size (from 6 to 21,432).

There are several noteworthy limitations of the current study. Measurement limitations included having the year of publication from low-, middle-, and high-income countries but with differences in time span between the years of publication for the three types of countries not being equivalent (i.e., from 1952 to 2015 for low-income countries, from 1950 to 2022 for middle-income countries, and from 1965 to 2022 for high-income countries). Accordingly, limiting variations on IQ change over time between types of countries may confound the current findings. In addition, despite the FE being investigated for several age groups, the IQ transformation does not account for age norming. Also, it is plausible that the FE is present for infants (Lynn, 2009), but not here studied. The current study investigated the FE using the SPM, which limits the age range to between 8 and 65 years old. Further studies may benefit from investigating the FE in younger children by using the IQ scores from the Color Progressive Matrices (CPM).

5. Conclusions

The current results provide further evidence for the FE in non-verbal intelligence in 71 countries over seven decades in a large-scale world-wide study. The FE appears to be stronger for middle-income countries, where it is likely that more critical improvements have affected IQ test performance. The FE was overall more prominent in younger generation cohorts, where perhaps environmental influences may have the most impact on mental functioning. The study illustrates how harvesting very large samples of people who have been tested on the more universally used SPM can further support the presence of the FE and help elucidate the factors that contribute to this phenomenon.

Acknowledgement

We are grateful to James Flynn for inspiring us to complete this study and note his modesty in reviewing our previous studies in which he cordially greeted us with the words ‘Flynn here’, his comments graciously helpful**.** We hope we have done his legacy justice in this current study**.** We are also grateful to Dr Lorna Morris and the anonymous reviewers for their insightful comments and careful reading of the manuscript.

Funding

The current investigation was financially supported by Cognitive Science and Innovation Research Unit **(**CSIRU**)**, College of Research Methodology and Cognitive Science, Burapha University, Thailand**.**

Declaration of Competing Interest

None**.**

References

Baker, D. P., Eslinger, P. J., Benavides, M., Peters, E., Dieckmann, N. F., & Leon, J. (2015). The cognitive impact of the education revolution: A possible cause of the Flynn effect on population IQ. *Intelligence, 49*, 144-158. https://doi.org/10.1016/j.intell.2015.01.003

Beller, J., Kuhlmann, B. G., Sperlich, S., & Geyer, S. (2022). Secular Improvements in Cognitive Aging: Contribution of Education, Health, and Routine Activities. *Journal of Aging and Health*, *34*(6–8), 807–817. https://doi.org/10.1177/08982643211065571

Berner, J., Comijs, H., Elmståhl, S., Welmer, A.-K., Sanmartin Berglund, J., Anderberg, P., & Deeg, D. (2019). Maintaining cognitive function with internet use: A two-country, six-year longitudinal study. *International Psychogeriatrics, 31*(7), 929-936. https://doi.org/10.1017/S1041610219000668

Booth, R. W., Sharma, D., & Leader, T. I. (2016). The age of anxiety? It depends where you look: changes in STAI trait anxiety, 1970–2010. *Social Psychiatry and Psychiatric Epidemiology, 51*(2), 193-202. https://doi.org/10.1007/s00127-015-1096-0

Bordone, V., Scherbov, S., & Steiber, N. (2015). Smarter every day: The deceleration of population ageing in terms of cognition. *Intelligence, 52*, 90-96. https://doi.org/10.1016/j.intell.2015.07.005

Bratsberg, B., & Rogeberg, O. (2018). Flynn effect and its reversal are both environmentally caused. *Proceedings of the National Academy of Sciences, 115*(26), 6674-6678. https://doi.org/doi:10.1073/pnas.1718793115

Brinch, C. N., & Galloway, T. A. (2012). Schooling in adolescence raises IQ scores. *Proceedings of the National Academy of Sciences, 109*(2), 425-430. https://doi.org/doi:10.1073/pnas.1106077109

Brouwers, S. A., Van de Vijver, F. J. R., & Van Hemert, D. A. (2009). Variation in Raven's progressive matrices scores across time and place. *Learning and Individual Differences, 19*(3), 330-338. https://doi.org/10.1016/j.lindif.2008.10.006

Clark, C. M., Lawlor-Savage, L., & Goghari, V. M. (2016). The Flynn effect: A quantitative commentary on modernity and human intelligence. *Measurement: Interdisciplinary Research and Perspectives, 14*(2), 39-53. https://doi.org/10.1080/15366367.2016.1156910

Colom, R., Andrés-Pueyo, A., & Juan-Espinosa, M. (1998). Generational IQ gains: Spanish data. *Personality and Individual Differences, 25*(5), 927-935. https://doi.org/10.1016/S0191-8869(98)00090-7

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*(2), 346-369. https://doi.org/10.1037/0033-295X.108.2.346

Donnelly, K., & Twenge, J. M. (2017). Masculine and feminine traits on the Bem sex-role inventory, 1993–2012: A cross-temporal meta-analysis. *Sex Roles, 76*(9), 556-565. https://doi.org/10.1007/s11199-016-0625-y

Dutton, E., van der Linden, D., & Lynn, R. (2016). The negative Flynn effect: A systematic literature review. *Intelligence, 59*, 163-169. https://doi.org/10.1016/j.intell.2016.10.002

Eppig, C., Fincher, C. L., & Thornhill, R. (2010). Parasite prevalence and the worldwide distribution of cognitive ability. *Proceedings of the Royal Society B: Biological Sciences, 277*(1701), 3801-3808. https://doi.org/10.1098/rspb.2010.0973

Feis, Y. F. (2010). Raven's progressive matrices. In C. S. Clauss-Ehlers (Ed.), *Encyclopedia of Cross-Cultural School Psychology* (pp. 787-787). Springer US. https://doi.org/10.1007/978-0-387-71799-9_344

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*(1), 29-51. https://doi.org/10.1037/0033-2909.95.1.29

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*(2), 171-191. https://doi.org/10.1037/0033-2909.101.2.171

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*, 171-191. https://doi.org/10.1037/0033-2909.101.2.171

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

Flynn, J. R., & Widaman, K. F. (2008). The Flynn Effect and the Shadow of the Past: Mental Retardation and the Indefensible and Indispensable Role of IQ. In *International Review of Research in Mental Retardation* (Vol. 35, pp. 121-149). Academic Press. https://doi.org/https://doi.org/10.1016/S0074-7750(07)35004-0

Galiciolli, M. E. A., Lima, L. S., da Costa, N. d. S., de Andrade, D. P., Irioda, A. C., & Oliveira, C. S. (2022). IQ alteration induced by lead in developed and underdeveloped/developing countries: A systematic review and a meta-analysis. *Environmental Pollution, 292*, 118316. https://doi.org/10.1016/j.envpol.2021.118316

Giangrande, E. J., Beam, C. R., Finkel, D., Davis, D. W., & Turkheimer, E. (2022). Genetically informed, multilevel analysis of the Flynn effect across four decades and three WISC versions. *Child Development, 93*(1), e47-e58. https:// doi.org/10.1111/cdev.13675

Greenfield, P. M. (2009). Technology and informal education: What is taught, what is learned. *Science, 323*(5910), 69-71. https://doi.org/10.1126/science.1167190

Ihle, A., Bavelier, D., Maurer, J., Oris, M., & Kliegel, M. (2020). Internet use in old age predicts smaller cognitive decline only in men. *Scientific Reports, 10*(1), 8969. https://doi.org/10.1038/s41598-020-65846-9

Jones, G., & Schneider, W. J. (2006). Intelligence, human capital, and economic growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *Journal of Economic Growth, 11*(1), 71-93. http://www.jstor.org/stable/40216088

Kanaya, T. (2016). Discussing the Flynn effect: From causes and interpretation to implications. *Measurement: Interdisciplinary Research and Perspectives, 14*(2), 67-69. https://doi.org/10.1080/15366367.2016.1171607

Kaufman, A. S., Zhou, X., Reynolds, M. R., Kaufman, N. L., Green, G. P., & Weiss, L. G. (2014). The possible societal impact of the decrease in U.S. blood lead levels on adult IQ. *Environmental Research, 132*, 413-420. https://doi.org/10.1016/j.envres.2014.04.015

Kim, Y. K., & Han, S. H. (2022). Internet use and cognitive functioning in later life: Focus on asymmetric effects and contextual factors. *Gerontologist, 62*(3), 425-435. https://doi.org/10.1093/geront/gnab149

Kokkinakis, A. V., Cowling, P. I., Drachen, A., & Wade, A. R. (2017). Exploring the relationship between video game expertise and fluid intelligence. *PLoS One, 12*(11), e0186621. https://doi.org/10.1371/journal.pone.0186621

Li, Y., Han, W.-J., & Hu, M. (2022). Does internet access make a difference for older adults' cognition in urban China? The moderating role of living arrangements. *Health & Social*

Care in the Community, 30(4), e909-e920. https://doi.org/https://doi.org/10.1111/hsc.13493

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis*. Sage Publications, Inc.

Lynn, R. (2009). What has caused the Flynn effect? Secular increases in the development quotients of infants. *Intelligence, 37*(1), 16-24 https://doi.org/10.1016/j.intell.2008.07.008

Martorell, R. (1998). Nutrition and the worldwide rise in IQ scores. In *The rising curve: Long-term gains in IQ and related measures.* (pp. 183-206). American Psychological Association. https://doi.org/10.1037/10270-006

Merten, N., Pinto, A. A., Paulsen, A. J., Chen, Y., Schubert, C. R., & Cruickshanks, K. J. (2022). Better cognitive function in younger generations - Insights from two cohort studies of middle-aged to older adults in Wisconsin. *Maturitas, 162*, 31-36. https://doi.org/10.1016/j.maturitas.2022.04.002

O'Keefe, P., & Rodgers, J. L. (2022). Home improvement: Evaluating secular changes in NLSY HOME-Cognitive Stimulation and Emotional Support scores. *Journal of Child and Family Studies, 31*(1), 1-16. https://doi.org/10.1007/s10826-021-02149-1

Oliver, M. B., & Hyde, J. S. (1993). Gender differences in sexuality: a meta-analysis. *Psychological Bulletin, 114*(1), 29-51. https://doi.org/10.1037/0033-2909.114.1.29

Orach, C. G. (2009). Health equity: challenges in low income countries. *African Health Sciences, 9 Suppl 2*(Suppl 2), S49-51.

Palma-Ruiz, J. M., Torres-Toukoumidis, A., González-Moreno, S. E., & Valles-Baca, H. G. (2022). An overview of the gaming industry across nations: using analytics with power

BI to forecast and identify key influencers. *Heliyon, 8*(2), e08959. https://doi.org/10.1016/j.heliyon.2022.e08959

Pietschnig, J. (2016). The Flynn effect: Technology may be part of it, but is most certainly not all of it. *Measurement: Interdisciplinary Research and Perspectives, 14*(2), 70-73. https://doi.org/10.1080/15366367.2016.1171612

Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains:A formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science, 10*(3), 282-306. https://doi.org/10.1177/1745691615577701

Pietschnig, J., Voracek, M., & Formann, A. K. (2010). Pervasiveness of the IQ rise: A cross-temporal meta-analysis. *PLOS ONE, 5*(12), e14406. https://doi.org/10.1371/journal.pone.0014406

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology, 41*(1), 1-48. https://doi.org/10.1006/cogp.1999.0735

Raven, J. C. (1936). *Mental tests used in genetic studies: The performance of related individuals on tests mainly educative and mainly reproductive* [University of London].

Reardon, S. (2014). Science in court: Smart enough to die? *Nature, 506*(7488), 284-286. https://doi.org/10.1038/506284a

Rindermann, H., & Becker, D. (2018). Flynn effect and economic growth: Do national increases in intelligence lead to increases in GDP? *Intelligence, 69*, 87-93. https://doi.org/10.1016/j.intell.2018.05.001

Rodgers, J. L. (1998). A critique of the Flynn effect: Massive IQ gains, methodological artifacts, or both? *Intelligence, 26*(4), 337-356. https://doi.org/10.1016/S0160-2896(99)00004-5

Rodgers, J. L. (2015). Methodological issues associated with studying the Flynn effect: Exploratory and confirmatory efforts in the past, present, and future. *Journal of Intelligence, 3*(4), 111-120. https://doi.org/10.3390/jintelligence3040111

Rönnlund, M., & Nilsson, L.-G. (2009). Flynn effects on sub-factors of episodic and semantic memory: Parallel gains over time and the same set of determining factors. *Neuropsychologia, 47*(11), 2174-2180. https://doi.org/10.1016/j.neuropsychologia.2008.11.007

Sauce, B., Liebherr, M., Judd, N., & Klingberg, T. (2022). The impact of digital media on children's intelligence while controlling for genetic differences in cognition and socioeconomic background. *Scientific Reports, 12*(1), 7720. https://doi.org/10.1038/s41598-022-11341-2

Schmidt, F. L., & Hunter, J. E. (2015). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage.

Schooler, C. (1998). Environmental complexity and the Flynn effect. In *The rising curve: Long-term gains in IQ and related measures.* (pp. 67-79). American Psychological Association. https://doi.org/10.1037/10270-002

Scopus. (2023). *Flynn effect*. Retrieved 14-Jan-2023 from www.scopus.com

Seferidi, P., Hone, T., Duran, A. C., Bernabe-Ortiz, A., & Millett, C. (2022). Global inequalities in the double burden of malnutrition and associations with globalisation: A multilevel analysis of demographic and health surveys from 55 low-income and middle-income countries, 1992-2018. *Lancet Glob Health, 10*(4), e482-e490. https://doi.org/10.1016/s2214-109x(21)00594-5

Shakeel, M. D., & Peterson, P. E. (2022). A half century of progress in US student achievement: Agency and Flynn effects, ethnic and SES differences. *Educational Psychology Review*. https://doi.org/10.1007/s10648-021-09657-y

Statista. (2022). *Percentage of global population accessing the internet from 2005 to 2021, by market maturity*. https://www.statista.com/statistics/209096/share-of-internet-users-in-the-total-world-population-since-2006/

Sternberg, R. J. (2019). *Human Intelligence: An Introduction*. Cambridge University Press.

te Nijenhuis, J., & van der Flier, H. (2013). Is the Flynn effect on *g*?: A meta-analysis. *Intelligence, 41*(6), 802-807. https://doi.org/10.1016/j.intell.2013.03.001

Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The Flynn effect: A meta-analysis. *Psychol Bull, 140*(5), 1332-1360. https://doi.org/10.1037/a0037173

Twenge, J. M. (2011). The duality of individualism: Attitudes toward women, generation me, and the method of cross-temporal meta-analysis. *Psychology of Women Quarterly, 35*(1), 193-196. https://doi.org/10.1177/0361684310395918

Twenge, J. M., Gentile, B., DeWall, C. N., Ma, D., Lacefield, K., & Schurtz, D. R. (2010). Birth cohort increases in psychopathology among young Americans, 1938–2007: A cross-temporal meta-analysis of the MMPI. *Clinical Psychology Review, 30*(2), 145-154. https://doi.org/10.1016/j.cpr.2009.10.005

Vainikainen, M.-P., & Hautamäki, J. (2022). Three studies on learning to learn in Finland: Anti-Flynn effects 2001–2017. *Scandinavian Journal of Educational Research, 66*(1), 43-58. https://doi.org/10.1080/00313831.2020.1833240

Voinea, C., Vică, C., Mihailov, E., & Savulescu, J. (2020). The Internet as cognitive enhancement. *Science and Engineering Ethics, 26*(4), 2345-2362. https://doi.org/10.1007/s11948-020-00210-8

Weber, D., Dekhtyar, S., & Herlitz, A. (2017). The Flynn effect in Europe – Effects of sex and region. *Intelligence, 60*, 39-45. https://doi.org/https://doi.org/10.1016/j.intell.2016.11.003

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), 1-20. https://doi.org/10.1016/j.intell.2009.05.002

Wongupparaj, P., Kumari, V., & Morris, R. G. (2015). A cross-temporal meta-analysis of Raven's progressive matrices: Age groups and developing versus developed countries. *Intelligence, 49*, 1-9. https://doi.org/10.1016/j.intell.2014.11.008

Wongupparaj, P., Wongupparaj, R., Kumari, V., & Morris, R. G. (2017). The Flynn effect for verbal and visuospatial short-term and working memory: A cross-temporal meta-analysis. *Intelligence, 64*, 71-80. https://doi.org/10.1016/j.intell.2017.07.006

Woodley, M. A. (2011). The Cognitive Differentiation-Integration Effort Hypothesis: A Synthesis between the Fitness Indicator and Life History Models of Human Intelligence. *Review of General Psychology, 15*(3), 228-245. https://doi.org/10.1037/a0024348