

The cognitive impact of the education revolution: A possible cause of the Flynn Effect on population IQ

David P. Baker^{a,b,*}, Paul J. Eslinger^{c,d}, Martin Benavides^{e,f}, Ellen Peters^g, Nathan F. Dieckmann^h, Juan Leon^e

^a Department of Sociology, Pennsylvania State University, United States

^b Department of Education Policy Studies, Pennsylvania State University, United States

^c Department of Neurology, Hershey Medical Center, Pennsylvania State University, United States

^d Department of Radiology, Hershey Medical Center, Pennsylvania State University, United States

^e Group for Analysis of Development (GRADE), Lima Peru

^f Department of Social Sciences, Pontificia Universidad Católica del Perú, Peru

^g Department of Psychology, Ohio State University, United States

^h Decision Research and Oregon Health & Science University, United States

ARTICLE INFO

Article history:

Received 25 March 2014

Received in revised form 10 December 2014

Accepted 14 January 2015

Available online xxxx

Keywords:

Education effects

Flynn Effect

IQ

ABSTRACT

The phenomenon of rising IQ scores in high-income nations over the 20th century, known as the Flynn Effect, indicates historical increase in mental abilities related to planning, organization, working memory, integration of experience, spatial reasoning, unique problem-solving, and skills for goal-directed behaviors. Given prior research on the impact of formal education on IQ, a three-tiered hypothesis positing that schooling, and its expansion and intensification over the education revolution, is one likely cause of the Flynn Effect is tested in three studies. First, a neuroimaging experiment with children finds that neuromaturation is shaped by common activities in school, such as numeracy, and share a common neural substrate with fluid IQ abilities. Second, a field study with adults from insolated agrarian communities finds that variable exposure to schooling is associated with related variation in the mental abilities. Third, a historical–institutional analysis of the cognitive requirements of American mathematics curriculum finds a growing cognitive demand for birth cohorts from later in the 20th century. These findings suggest a consilience of evidence about the impact of mass education on the Flynn Effect and are discussed in light of the g-factor paradigm, cognition, and the Bell Curve debate.

© 2015 Published by Elsevier Inc.

The development and spread of mass schooling over the past 150 years, known as the education revolution, represents a major social trend with nearly full enrollments attained in primary and secondary schools first in wealthier nations and since the middle of the twentieth century in most nations (e.g., Baker, 2014). The United States led the way in developing formal education intended for all children and youth by expanding gross enrollment rates from about 50% early in the

20th century to almost 90% by 1960, with mean education attainment levels increasing from 6.5 to 12 years over the same period (U.S. Department of Education, 1993). In addition to expanding exposure to formal education, the education revolution has given rise to a culture of academic achievement from early childhood education up through university graduate programs (e.g., Martinez, 2000; Nisbett, 2009). In distinct contrast to traditional societies, cognitive skills such as literacy, numeracy, and mastery of other academic subjects are assumed necessary for a successful adult life for all individuals. Furthermore, success in school and social status attainment are increasingly linked together across successive generations, thus

* Corresponding author at: Dept of Education Policy Studies, Penn State University, 310a Rackely Bld. University Park PA 16801, United States.

E-mail address: dpb4@psu.edu (D.P. Baker).

the motivation to attend and achieve in school reinforce one another (e.g., Bills, 2004; Torche, 2011).

Over the course of the education revolution, evidence exists that schooling has become more “cognitized,” in that higher order thinking and reasoning capabilities such as unique problem solving, flexible thinking, abstraction, informed interpretative skills, relational reasoning, generating new ideas and critique are increasingly assumed essential for successful academic achievement (e.g., Genovese, 2002; Martinez, 2000). And prior neurological and cross-cultural research point to the possibility that the learning of basic literacy, numeracy, and other academic subjects, even under rudimentary conditions for a limited time, leads to neural and cognitive enhancements (Baker, Salinas, & Eslinger, 2012). But the impact of the education revolution on changes in the cognitive capacity across generations is not clear. Schooling is often presumed to impart facts, information, and specific skills more than it is thought to influence general cognitive ability. The Flynn Effect – rising IQ scores in developed nations over the 20th century – offers a case for investigation of the impact of the education revolution’s positive impact on the cognitive capacity of populations.

A three-tiered hypothesis is developed positing that schooling, with its expansion and intensification over the education revolution, is one likely cause of the Flynn Effect. The parts of the hypothesis are tested in three studies: 1) a neuroimaging experiment, 2) a field study of exposure to schooling and cognitive skill enhancement, and 3) a historical-institutional analysis of the cognitive demands of schooling. The implications of the results for the study of intelligence, its consequences, and the Bell Curve argument are discussed.

1. Rising population cognitive performance and exposure to education

Mean IQ test scores of cohorts of American adults increased by approximately 25 points over the last 90 years, a period during which successive cohorts of children and youth were exposed to more formal education as shown in Fig. 1. For Japan, South Korea, and nations of Western Europe, where there is a

history of large-scale IQ testing, similar Flynn Effects and increases in educational attainment have been reported (Flynn, 1984, 1987; Lynn, 2009a; Lynn & Meisenberg, 2010; te Nijenhuis, Cho, Murphy, & Lee, 2012; UNESCO, 2002). The gains in crystallized intelligence reflected by the Wechsler Adult Intelligence Scale (WAIS) are surpassed by even steeper historical gains in fluid intelligence, assessed for example with Raven’s Progressive Matrices Test. Fluid intelligence includes the effective use of cognitive executive functions that provide mental resources for planning, organization, working memory, integration of experience, spatial reasoning, unique problem-solving, and skills for goal-directed behavior related to reasoning ability as applied in novel contexts. These “cognitive executive functions” (hereafter CEFs) are widely hypothesized to be the foundation for domain-general intelligence (g) that IQ tests attempt to measure and are considered shared cognitive resources applicable across multiple specific content areas and situations (e.g., Nisbett et al., 2012). The focus here, fluid intelligence gains relying heavily on CEFs are approximately double those of crystallized cognitive ability gains, and are generally considered to have occurred too rapidly to be attributable to genetic selection (Flynn, 2009). Lastly, the Flynn Effect has been judged to be substantial and is corroborated by studies from early in the 20th century (e.g. Lynn, 2013; Runquist, 1936).

Since Spearman (1923) the reigning theoretical paradigm posits that intelligence is an underlying trait g , assumed to be an expression of mental ability, significantly inherited, and highly stable across successive cohorts. And although modern IQ tests experiencing the Flynn Effect were originally developed to predict educability of individuals, the g -factor paradigm assumes that such tests are accurate indicators of an individual’s underlying domain-general intelligence. Supporting this paradigm are well-known studies of monozygotic twins raised separately that report IQ measures of intelligence reliably factoring into one latent general factor like g , which is associated with real world competence and is substantially inherited at levels between 0.50 and 0.75 (e.g., Gottfredson, 1997; Jensen, 1998).

In contrast, 3–7 point average gain each decade across same-aged test-takers on IQ tests (that have a mean of 100 points and

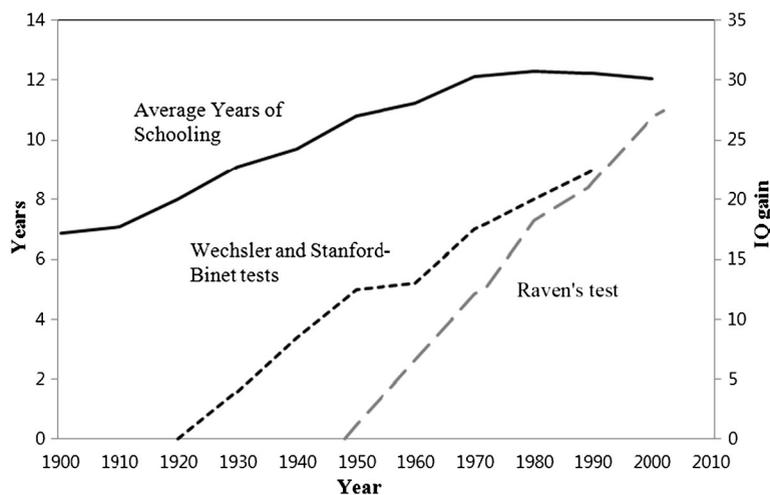


Fig. 1. Growth in educational attainment and adult IQ test score gains in the United States during the 20th century.

Sources: Average years of schooling completed by age 20: IPUMS; Wechsler and Stanford-Binet test (whites: Horgan, 1995); Raven’s test IQ: Flynn (2009).

standard deviation of 15 points) poses a significant challenge to the *g*-factor paradigm. For example, if the range of *g* is time invariant, a 1.5 standard deviation increase in average scores across three generations would indicate that most individuals born in the first third of the 20th century were cognitively dull and unable, for example, to understand the rules of popular games such as baseball, whereas over half of those born since 1970 would be classified as cognitively superior with many in the genius range. That this is clearly not true poses a further challenge to the assumptions of the *g*-factor paradigm in light of the Flynn gains in CEFs (Neisser, 1998).

The Flynn Effect begs the question of what environmental factor or factors could cause these gains. Proposed hypotheses exist on improving nutrition, rising television viewing, electronic game consumption, growing social complexity, and mass media exposure of test items, but no existing proposals meet the necessary conditions to be responsible for change across the entire time period (Neisser, 1998). While it is plausible that a variety of causal factors functioning at different historical points could explain the effect, a consistent factor with the same temporal increase as gains in CEFs over the 20th century is important to test and could be part of a parsimonious explanation. In particular, a hypothesis has three necessary conditions to be considered as a cause of the Flynn Effect over the entire period (Blair, Gamson, Thorne, & Baker, 2005):

- 1) Neurologically, the environmental factor must be shown to have an ontogenetic effect on individual cognition.
- 2) Cognitively, variation in exposure to the environmental factor must be shown to cause between-individual variation in cognitive skills like those tested on IQ measures.
- 3) Demographically and institutionally, exposure to the environmental factor must be:
 - a. present in the population over the entire period of rising population IQ scores;
 - b. shown to increase monotonically across successive birth cohorts over the period up to full population saturation;
 - c. the intensity of the causal mechanism must increase as exposure reaches nearly all individuals in later cohorts.

Many suggested causes, such as growing social complexity, lack evidence to suffice condition 1. Significant television consumption, for example, has not been shown to meet conditions 1 or 2, and it does not meet 3a. Use of electronic video games has been associated with enhancement of one type of fluid cognitive skill (condition 2), but it does not meet conditions 3a and 3b (e.g., Okagaki & Frensch, 1994). Similarly, there is no evidence that IQ test items have been widely reproduced in mass media, or that possible leaks increased monotonically across birth cohorts, although there is new evidence about test-taking strategies (3b) (Hayes, Petrov, & Sederberg, 2015). Lastly, the most researched hypothesis to date is that rising population nutrition is one contributor to the Flynn Effect. Although evidence exists that increased nutrition has been associated with larger average head sizes over cohorts, and meets conditions 1 and 2 (at least from a deficiency perspective), there is also counter evidence (e.g., Flynn, 2009; Lynn, 1989; Martorell, 1998).

The historical demography of the education revolution does meet conditions 3a and 3b. Despite earlier conjecture that formal schooling could be a cause of the Flynn Effect, this hypothesis was not been tested because of widely-held misconceptions about the institutional development of schooling and its

potential influence on cognition (e.g., Williams, 1998). A good example is the untested presupposition that curriculum has become “dumbed down” over the last century, thus failing condition 3c (this is addressed in the third study below). But by far the most limiting misconception is that schooling cannot meet conditions 1 and 2. For example, much-reported modest between-school and larger between-family effects on early school achievement are often misinterpreted to mean that the effects of exposure to schooling must be too modest to meet condition 1. However, there is a growing empirical literature indicating support for both conditions.

Tested here is a schooling-cognitive enhancement hypothesis that meets these causal conditions and is comprised of three sub-hypotheses based on prior research and tested in corresponding studies below:

H1. Exposure to common schooling activities will activate the neural substrate responsible for the development of CEFs.

H2. Variation in exposure to schooling causes between-individual differences in expressed CEFs.

H3. Educational expansion monotonically increased across birth cohorts, and the content of schooling has intensified in material requiring greater use of CEFs in later decades of the 20th century.

2. Three tests of the schooling-cognition enhancement hypothesis

2.1. Study 1. Neuroimaging experiment: conjoint fMRI analysis of CEFs and calculation schooling tasks

Establishing schooling-related activation of the neural substrate involved in CEF performance is a necessary first step to test H1 and satisfy causal condition 1. Functional magnetic resonance imaging (fMRI) and analysis of its blood oxygenation level dependent (BOLD) changes yield data about regional brain activation during performance of specific tasks. The ontogenetic-causal logic behind fMRI evidence is supported by a large set of empirical findings concluding that the neurodevelopment of mammals is caused by interplay between activity-dependent neural plasticity and neurobiological maturation processes (e.g., Quartz & Sejnowski, 1997). fMRI studies aim to provide evidence of specific activity-dependent plasticity affecting neurocognitive development as a function of environmental factors; and significant effects are reported among musicians, taxi drivers, jugglers, exercisers, whole body balancers, and medical students, who all have experienced specified, prolonged environmental stimulation and engagement in order to achieve defined expertise in their respective pursuits (see review in Baker et al., 2012).

In terms of school-based learning tasks and regional brain activation, anatomical differences exist due to developmental and adult acquisition of literacy skills. In particular, such training (compared to no-additional training) in adults resulted in changes in brain anatomy including an enlarged splenium of the corpus callosum (where visual pathways cross between the dominant language and nondominant hemispheres) and increased tissue volumes in multiple cortical regions (i.e., the inferior parietal, dorsal occipital and middle temporal cortices,

the left supramarginal gyrus and superior temporal cortex) associated with academic skills of oral language comprehension, reading, writing, and numeracy. Conversely, the white matter of the corpus callosum in auditory and language processing areas is comparatively thinner in illiterates compared to literate populations (Ardila et al., 2010; Carreiras et al., 2009; Castro-Caldas et al., 1999). Hence, it is feasible that specific exposure to intense academic training will modify brain size and organization. It has also been demonstrated that adults' CEFs can be enhanced by exposure to specific cognitive training (Willis et al., 2006). There is other evidence that school-related tasks may contribute to neural plasticity of CEFs. For example, results exist of consistent developmental activations in the frontal and parietal lobe regions in children undertaking various kinds of symbolic and non-symbolic calculation tasks (Kaufmann, Wood, Rubinstein, & Henik, 2011; Rivera, Reiss, Eckert, & Menon, 2005). These frontal and parietal regions are also implicated in higher level integrative processing characteristic of CEFs (Naghavi & Nyberg, 2005) and are activated in children and adolescents responding to a novel problem solving task requiring CEFs (Eslinger et al., 2009). And early damage to frontal cortical regions has been shown to impair development of CEFs (Eslinger, Flaherty-Craig, & Benton, 2004). All of these studies suggest that CEFs are shaped by the interaction between genetic and environmental factors.

No study to date has analyzed two common academic tasks that are thought to involve CEFs — i.e., mathematical calculation and relational reasoning — with respect to their shared skills and neural demands. To address this issue, a within-subject experiment was undertaken to test the hypothesis that common regional brain activations occur while children solve calculation problems from school math curricula and relational reasoning problems characteristic of CEF skills. Support for the hypothesis that exposure to common schooling activities will activate the neural substrate responsible for CEF development requires that anatomically similar, known as conjoint, activations are associated with common calculation skills and with higher-order CEF skills. Such math skills are important to study because they comprise 60–80% of the content of school math textbooks (see Study 3) and could contribute to activity-dependent development in the cortical regions associated with CEFs. fMRI conjunction analysis was employed to provide this test.

2.1.1. Methods

2.1.1.1. Subjects. Participants (N = 16; 8–19 years of age; 69% males) were recruited from a medium-sized town in the U.S. for the fMRI study. Participants were included only if they had no history of medical, neurological or psychiatric illness, learning disability, or current medication usage. They were screened with standardized tests of general intellect, academic achievement, and CEF, with all scores varying within the normal range. Average age was 12.81 years (s.d. 3.25) and average education level was 6.38 (s.d. 3.20). Average scores on the Wide Range Achievement Test-3 (Wilkinson, 1993) were 109.0 (s.d. 9.43) for reading and 107.24 (s.d. 11.45) for arithmetic.

2.1.1.2. fMRI procedures. Participants were familiarized with the 3 T magnetic field environment, the in-magnet visual display, and the button-press-response equipment for the tasks employed in this study (see Appendix A for details of fMRI

procedures). Prior to going into the magnet, participants reviewed all instruction slides and practiced sample problems on a desktop computer; all participants completed the studies without difficulty. Conjunction analysis was performed to examine the brain activation patterns associated with the CEF relational reasoning task and the calculation task (Price & Friston, 1997). This analysis identifies if there is any common area(s) of brain activation for this pair of experimental tasks.

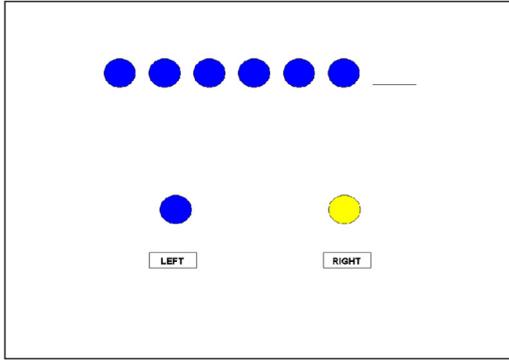
2.1.1.3. Tasks and design. Each task type required subjects to complete a baseline condition so that brain activations specifically associated with relational reasoning and calculation could be isolated from general cognitive activation effects. Tasks were presented in three runs of baseline stimuli during a 78-second baseline condition and three runs of experimental stimuli during a 78-second experimental condition. There was also a 15-second rest condition (no stimuli) interleaved. The baseline and experimental condition blocks each consisted of 16 stimuli with a 4-second exposure time and a 0.875 second inter-stimulus interval blank screen (total time = 78 s). For the relational reasoning task, baseline stimuli consisted of geometric designs of a single dimension that required a simple perceptual matching response choice (see Fig. 2A for example). The experimental stimuli were composed of colored geometric designs that varied along multiple dimensions of color and shape, requiring relational reasoning to identify the correct response from two choices. Subjects viewed reminder instructions before each task. The timing and switching of visual stimuli were automatically controlled by TTL signals incorporated in the pulse-timing program. For calculation tasks, participants responded to a simple counting baseline task in which they were shown a brief orienting question and a set of coins, requiring a yes or no button press response. The experimental condition presented coin calculation problems in which participants were shown a brief orienting question and a set of coins, requiring a numeric calculation and yes or no button press response (see Fig. 2B). The baseline and experimental blocks each consisted of 16 sets of stimuli, each with a 3-second exposure, and a 1-second blank screen inter-stimulus interval.

2.1.2. Results

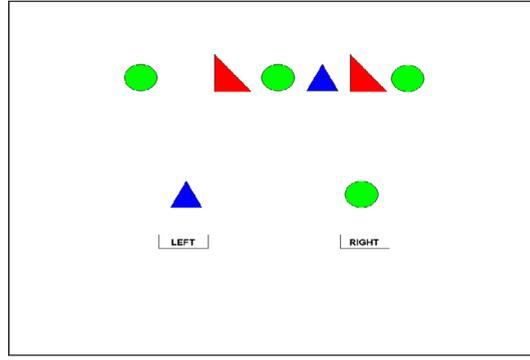
Data from each experimental task were statistically compared to its respective baseline condition to identify the specific activations associated with relational reasoning and symbolic calculation. With these brain activity maps, neural responses to the relational reasoning and calculation tasks were then compared. Conjunction analysis revealed that common activations were detected in the parietal, frontal, and occipital lobes (see Fig. 3 for summary and Table 1). Of particular interest, significant conjoint activations occurred in the left and right superior parietal cortices (including the intraparietal sulcus area that has been implicated in primary numeric processes, Dehaene, 1996) and in surrounding cortices implicated in visuospatial relational reasoning, working memory and calculation, and the left and right dorsolateral prefrontal cortices implicated in relational reasoning processing and working memory (Eslinger et al., 2009).

As with prior research on relational reasoning, frontal activations in our study occurred in the region of the middle frontal gyrus whereas superior parietal activations occurred in the region of the intraparietal sulcus (e.g., Kroger et al., 2002).

A. Relational Reasoning Task (“what shape/color completes the pattern”)



BASELINE CONDITION

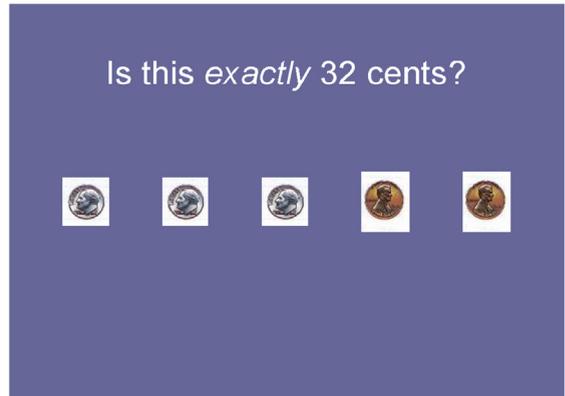


EXPERIMENTAL CONDITION

B. Calculation Task



BASELINE CONDITION



EXPERIMENTAL CONDITION

Fig. 2. Sample stimuli from the fMRI experiment. A. Relational reasoning task (“what shape/color completes the pattern”). B. Calculation task.

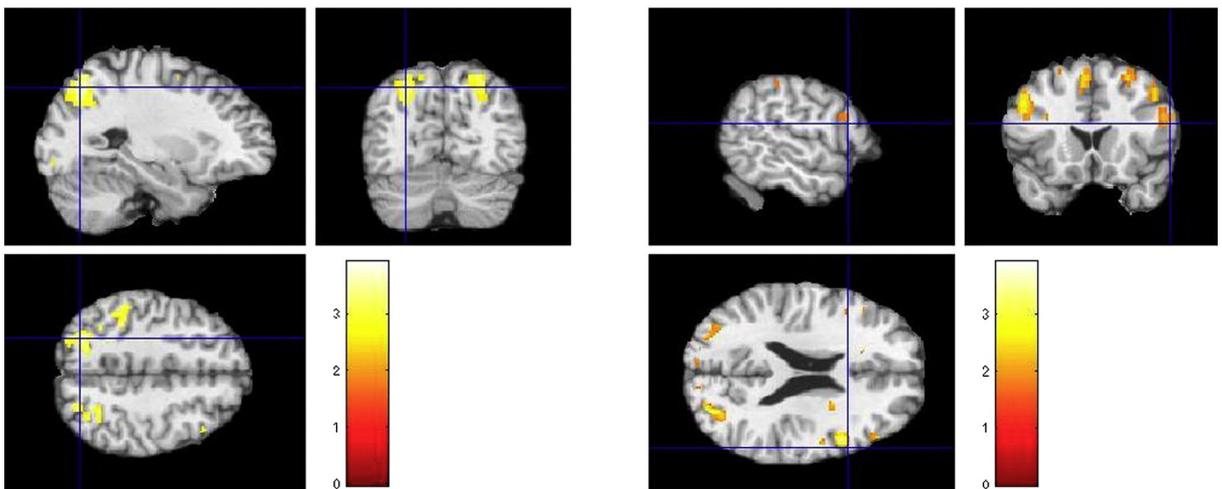


Fig. 3. Activated brain regions common to both relational reasoning and calculation tasks in a healthy developmental sample. Significant neural recruitment was detected in both superior parietal (seen in multiple planes in the left figures) and prefrontal (seen in multiple planes in the right figures) regions.

Table 1

Regions of common neuroanatomical activity associated with both relational calculation and reasoning tasks in a typically developing sample of children and adolescents.

Anatomical area	MNI coordinates			Activation size (pixels)	t value ^a
	X	Y	Z		
Parietal					
Left superior parietal (intraparietal sulcus region)	−25	−68	50	262	3.92
Left inferior parietal (angular gyrus region)	−38	−40	45	48	3.23
Right superior parietal (intraparietal sulcus region)	27	−72	50	174	3.37
Prefrontal					
Left prefrontal dorsolateral	−50	16	35	13	2.73
Right prefrontal ventrolateral	49	11	25	48	3.79
Right prefrontal dorsolateral	25	−4	55	21	3.49
Occipital					
Left occipital	−31	−86	−5	23	2.77
Right occipital	32	−79	10	106	3.36

Note. Standard MNI (Montreal Neurological Institute) coordinates represent the specific location of significant blood oxygen level dependent (BOLD) activity within 3-D brain mapping space. The activation size refers to the number of pixels comprising the anatomical cluster of each BOLD response (minimum threshold 10).

^a All t values < .05.

The common activations observed during relational reasoning and calculation show some overlap with recruited neural activity associated with spatially directed attention, and shifts in attention. These processes likely aid accuracy and efficiency in the experimental measures. The activations in the prefrontal cortex involved dorsolateral regions that have also been linked to spatial and non-spatial working memory critical to high demands for *integrative* processing (e.g., Deco, Rolls, & Horwitz, 2004). Such processing is thought to be *domain general* and not limited to specific materials, thus raising the possibility of a cognitive resource that can be applied to multiple content areas. Taken together, these results support H1 that common calculation tasks used in school and relational reasoning tasks lead to activation of CEF regions. They therefore satisfy a major condition for exposure to schooling as activating the neural substrate responsible for CEF development (causal condition 1).

2.2. Study 2. Field study and test of H2: variable exposure to schooling and enhancement of CEFs

The demography of the education revolution makes it difficult to directly compare cognitive functioning of normally developed individuals who are similar except for exposure to schooling. Children at different levels of schooling can serve as a proxy for exposure to schooling, but the confounding effects of neurological and cognitive maturation are difficult to overcome within such a design. Nevertheless recent studies of students report that exposure to schooling is associated with IQ test increases three to four times more than physical maturation (Cliffordson & Gustafsson, 2008; Clouston et al., 2012; Hatch, Feinstein, Link, Wadsworth, & Richards, 2007). These results are consistent with an earlier meta-analysis and reports of a declining cognitive performance during summer absence from school, and early field studies with adults (Ceci, 1991; Downey, von Hippel, & Broh, 2004; Luria, 1976). Certainly, children enter

school with differing levels of genetically endowed potential for intelligence and differing influences of early parenting. But, whatever the level of entrance characteristics, they are then immersed in a structured curriculum and sustained learning environment that prioritizes cognitive abilities related to CEFs (Flynn, 2012; Genovese, 2002).

While these prior studies fail to reject hypotheses similar to H2 (that variation in exposure to schooling causes between-individual differences in expressed CEFs), they suffer from insufficient controls for possible confounding variables. For example, field studies from the 1930s lacked controls on possible spurious factors and had crude instrumentation. The best test of H2 to establish causal condition 2 would be an experiment with random assignment of children to varying levels of exposure to schooling, including no schooling, and undertaking CEF measurement into adulthood. For obvious reasons, such an experiment is impossible, but there are populations living in relatively remote areas with adults who experienced significantly different exposures to schooling but whose post-schooling work and economic status are similar so that they can approximate an experiment on the cognitive effects of schooling. Although the variation in school attainment is not random in such populations, it is unlikely associated with actual or perceived differences in the individuals' pre-school IQ or aptitude for academic study (see analysis of possible selection bias below). Small, relatively isolated, agrarian Quechua-speaking communities in Peru's Andean highlands, in which schooling was rarely fully accessible or compulsory until recently, provides a suitable population for a superior field study of the effects of varying amounts of schooling on enhancement of CEFs.

2.2.1. Methods

2.2.1.1. Subjects. A sample of 247 adults aged 30 to 60 years old with a wide range of school attainment from unschooled to advanced schooling (mean 7.4 grade in school, s.d. 5, with 15% having no formal schooling) were selected based on a two-stage stratified sampling procedure. First, using Peru's National Census 2007, a list was developed of all small traditional agrarian communities within the Carhuaz district in the Ancash region and 14 communities were selected based on availability of maximum within-community variation in educational attainment among inhabitants and homogeneity of occupational structure (50% were subsistence-level farmers, and the remainder were employed in jobs in the local agrarian economy). Second, a door-to-door survey was conducted to recruit subjects, stratified by education attainment, with good general health and no past neurological trauma or evidence of brain damage due to alcohol abuse.

2.2.1.2. Dependent variables. Five standard, psychometrically developed, instruments measuring overlapping components of CEFs were used: 1) verbal association fluency test; 2) backward digits task of working memory capacity (Wechsler, 1981); 3) Delis–Kaplan Executive-Function System Tower test of planning, strategy, working memory, and attention shifting abilities (Delis, Kaplan, & Kramer, 2001); 4) Woodcock–Johnson – III Calculation test (first 30 items) of mathematical calculation, reasoning, and number sense (McGrew, Wood, & Mather, 2001); and 5) Raven's Colored Progressive Matrices test of non-verbal reasoning about complexity (Raven, Raven, & Court, 2000).

2.2.1.3. Independent variable. Exposure to schooling was measured through participant reports of highest grade of schooling completed.

2.2.1.4. Conditioning variables. The Peabody Picture of Vocabulary Test (PPVT), measuring crystallized verbal abilities and long-term memory, was administered to control for aspects of intelligence less reliant on CEF skills. Additionally, conditioning variables included age, gender, household and agricultural assets, religion, occupation, village size, and working history. For analysis of selection bias, participants were interviewed about the causes of their education attainment, with a special focus on reasons for under-schooled and no formal schooling.

2.2.1.5. Instrument preparation and interviewing. All cognitive instruments and questionnaire items were forward and backward translated from English to Spanish and then to the local dialect of Quechua. Extensive pilot testing and adaptation of verbal instructions, subject understanding of items and tasks, cultural appropriateness of instruments, and interviewing procedures were undertaken. Quechua-speaking interviewers (often local teachers) were trained and tested on administration of all instruments, and interviews occurred in subjects' homes or at villages' school buildings. Subjects were given a modest amount of household goods for their participation and their communities' schools were given educational materials.

2.2.2. Estimation

Models of the effect of exposure to schooling on adult CEFs were estimated within a Structural Equation Modeling (SEM) statistical approach (estimated in MPLUS 4.2; e.g., Kline, 1998). The model included two parts: 1) a measurement model estimating a latent variable of CEFs in a confirmatory factor analysis of the five CEF instruments and 2) a structural model of the effect of schooling (based on highest grade completed) on the latent construct of CEFs, conditioning on the variables described above.¹

1

Verbal associative fluency = $\lambda_1\eta_1 + \mu_1$

Working Memory = $\lambda_2\eta_1 + \mu_2$

Plan-strategy = $\lambda_3\eta_1 + \mu_3$

Numeracy Cal. = $\lambda_4\eta_1 + \mu_4$

Decision-making Abil. = $\lambda_5\eta_1 + \mu_5$

Non-verbal reasoning = $\lambda_6\eta_1 + \mu_6$

where:

λ_i factor loadings $i = 1, 2, 3, 4, 5, 6$
 μ_i measurement errors $i = 1, 2, 3, 4, 5, 6$

$\eta_1 = \beta_0 + \beta_1X + \beta_2W + \varepsilon_1$ (2)

where:

η_1 CEFS – latent variable with six indicators.
 X conditioning variables. Wyears of schooling.
 β Structural parameters.
 ε Random errors.

2.2.3. Results

Panel A of Fig. 4 displays the estimation of education's influence on the latent variable of CEFs without any controls and indicates that an increase of one standard deviation in highest grade of schooling increases CEFS ($\beta = 0.85, p = .001$). As shown in Panel B, after conditioning on demographic variables and an indicator of crystallized intelligence, the original schooling effect is moderated, but still remains positively and substantially related to CEFS ($\beta = 0.63, p = .001$).² Model 2 indices revealed an acceptable fit between the model and the observed data (RMSEA = .05; CFI = .95; TLI = .93).

2.2.3.1. Selection bias and endogeneity. The results above reflect a design using extensive conditioning including a concurrent measure of crystallized intelligence, and fail to reject H2, and support the proposition that formal education is a cause of increasing CEFs. However, without random assignment to educational attainment, the possibility remains that individuals with higher pre-school CEFs were selected to attend and to stay in school for longer. In addition, other sources of endogeneity – correlated, unobserved factors relegated to the error term – could mask the true direction of the effect and possible third factors could cause both greater schooling and enhanced CEF.

Three possible sources of selection bias and endogeneity were analyzed and found unlikely to be responsible for the results. The first source is that more cognitively talented individuals could have been selected to attend schooling. However, the reasons individuals gave for their school attendance were limited to non-cognitive factors, such as access to schooling, financial circumstances related to school fees, gender preference, or death of a parent or sibling. Further, since school attendance was not normative in these communities when subjects were young, social stigma for low or non-attendance is unlikely and thus there is little reason for subjects to fabricate excuses. Coding each participant's attendance reason and re-estimating the model found no change in the schooling to CEF coefficient, either in sign or in relative size. Second, a possible source of both selection bias and endogeneity is the effect of an individual's family (prior to that individual's schooling) on CEFs and education attainment. However, given that parental education is a primary cause of both preschool cognitive enhancement and accurate perceptions of young children's cognitive aptitude (Schaub, 2010) and that most participants' parents were unschooled (schooling in these communities was very rare when the parents of the subjects were school aged), it is likely that only limited and invariant parental influences existed in our participant sample. Third, it could be that more educated individuals showed greater post-schooling effects on CEF, but the complexity of jobs participants held varied little, and conditioning statistically for job type did not alter the schooling to CEF coefficient. Lastly, these results converge with similar studies in rural Ghana and

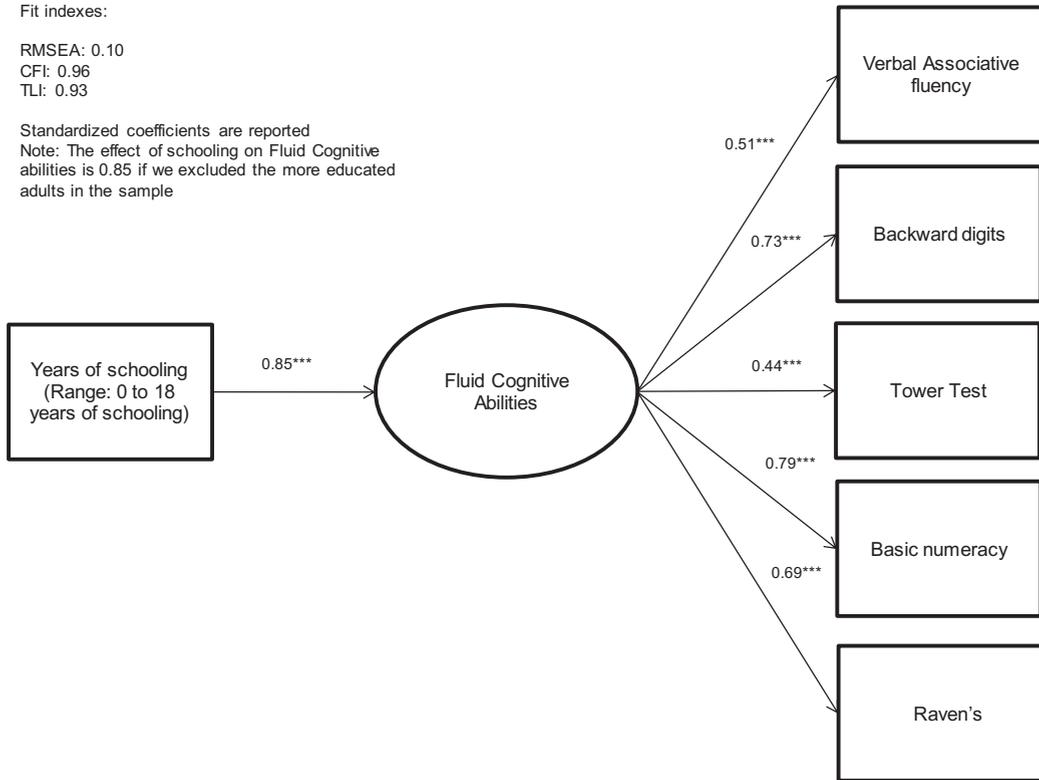
² Because a few individuals in our sample had high exposure to secondary schooling and whose post-schooling cognitive experiences could be more enriching than others, the model was re-estimated without these subjects and the exposure to schooling was still positively related to CEFS ($\beta = 0.65, p = .001$). This supplemental analysis and a matrix of all estimated coefficients for all estimates of the model described here are available from the authors upon request.

A

Fit indexes:

RMSEA: 0.10
CFI: 0.96
TLI: 0.93

Standardized coefficients are reported
Note: The effect of schooling on Fluid Cognitive abilities is 0.85 if we excluded the more educated adults in the sample



B

Fit indexes:

RMSEA: 0.06
CFI: 0.94
TLI: 0.93

Standardized coefficients are reported
Note: The effect of schooling on Fluid Cognitive abilities is 0.68 if we excluded the more educated adults in the sample

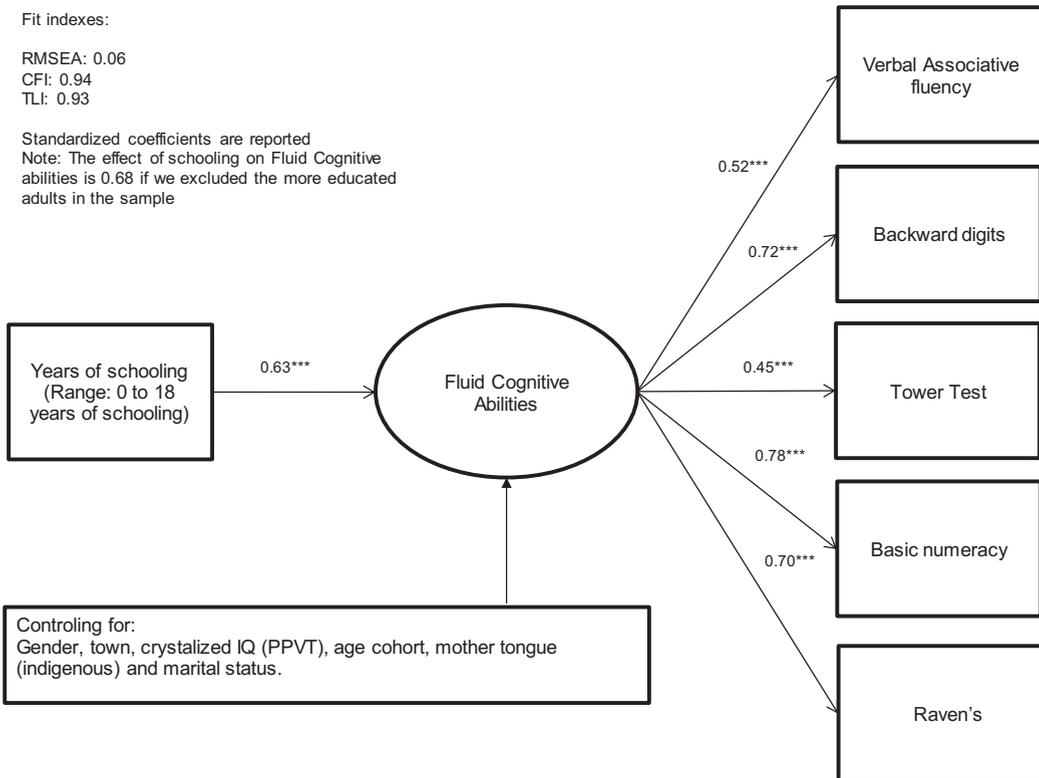


Fig. 4. Structural equation model of association between Peruvian subjects years of schooling and cognitive executive functioning.

Malawi (Peters, Baker, Dieckmann, Leon, & Collins, 2010; Van de Vijver & Brouwers, 2009).

2.3. Study 3. Historical–institutional analysis and test of H3: the cognitive demand of primary school curricula from 1930 to 2000

The well-documented demographic record of American school expansion and growing generational norms of educational attainment confirm that schooling as a potential environmental cause of population IQ increase has been present for the entire history of the trend and has monotonically grown in exposure across birth cohorts. Schooling thus meets causal conditions 3a and 3b (see Fig. 1). But, because the pace of this trend exposed nearly full generations of American children and youth to primary schooling by mid-century and to significant secondary schooling by the 1960s, testing H3 (causal condition 3c) is necessary to examine if the proposed causal agent in schooling increased its effect over later decades of the 20th century. In other words, did the core cognitive environment of schooling increase with each new generation of students post full demographic exposure to schooling circa 1960?

There are a number of well-known trends suggesting that enrollment expansion was accompanied by a growing intensity of schooling. The length of the American school year grew from about 6.5 months just before the beginning of the 20th century to the current nine months of schooling, or over a one third increase in instruction time (U.S. Department of Education, 2006, 2008). Also, the average student/teacher ratio has fallen steadily from the end of the 19th century so that by 2007, there were an estimated 15.4 public school pupils per teacher nationwide. In constant dollars, the average total expenditure on public school students grew from \$355 in 1919 to \$9518 per student in fiscal year 2004.

These indicators are suggestive, but they do not directly measure change in the core cognitive environment nor did they happen primarily as full exposure to primary and secondary school was occurring in the latter part of the 20th century. A more direct test of H3 (that educational expansion monotonically increased across birth cohorts, and the content of schooling required greater use of CEFs in later decades of the 20th century) would be an analysis of historical changes in the curriculum of mass schooling. A recent extensive content analysis of primary schooling mathematics textbooks from 1900 to 2000 provides the necessary data for this subject area. Textbooks are the most reliable and comprehensive surviving historical record of the U.S. elementary mathematics curriculum and can serve as an accurate indicator of what was taught in school (e.g., Baker et al., 2010). Of course, teachers may modify aspects of the content and suggested pedagogy in textbooks; they can augment them with other materials, selectively cover the content, and perhaps never use the purchased text. However, for the bulk of the nation's teachers during each historical period, it is a reasonable assumption that popular texts are a general indicator of the main trends of the elementary school mathematics curriculum and, hence, provide archival evidence of the nature of the core cognitive environment of schooling. To not reject H3, therefore, it should be the case that the cognitive demand of primary school mathematics curricular material grew in the latter half of the century, after full demographic exposure to schooling had been

reached. To make the analysis as conservative as possible, we focus on arithmetic, the most basic of mathematics topics and among textbooks for first and fourth grades.

2.3.1. Methods

2.3.1.1. Textbook sampling and analysis. The curricular data come from a content analysis of over 28,000 pages from a sampling of 141 widely-used school mathematics textbooks for grades 1–6 published by 10 different publishing companies. The entire corpus of sampled material was coded for its curricular content across various topical dimensions. Baker et al. (2010) provided a detailed description of textbook sampling, periodization, general content coding, coding reliability, and results of general curricular content trends over the 20th century. Coders were trained and reliability checks were completed to assure consistency.³ The analysis here was conducted on a sub-sample of these data consisting of a 20% random sampling of first and fourth grade textbook pages (1600 pages) that included coverage of strategies to solve problems using basic arithmetical operations beginning in 1932 as the era just prior to full exposure to primary education in the 1940s and 50s.⁴

2.3.1.2. Measurement of cognitive demand. Learning multiple solution strategies for problems and using other specific strategies that require effortful cognitive skills have been shown to be closely associated with demand on CEFs (e.g., Bull & Scerif, 2001). Therefore, the level of cognitive demand of the sampled textbook material was assessed in three ways. First, the frequency of each kind of problem-solving strategy was recorded across the sampled pages: more strategies increase the cognitive demand of students to learn, select, and evaluate each approach. Second, then each kind of problem-solving strategy was assessed as to the level of conceptual reasoning about mathematics required of the student to use the strategy effectively. A scale for coding was used from no mathematical reasoning (e.g., memorization of addition tables, counting objects) to a mid-point of moderate conceptual reasoning (e.g., using addition as union of sets) to an end-point of significant conceptual reasoning. The operationalization of significant conceptual reasoning included having to do operations that maximized student's existing knowledge or that decomposed more difficult problems into familiar ones (e.g., commutative property of addition for real numbers). The more problems a textbook had with higher scores on the scale, the more cognitive demand it required. Lastly, the frequency of four specific strategies that demanded intense use of CEFs was also calculated: 1) explicit reasoning about mathematics with or without formal calculations; 2) estimation of problem solutions; 3) mental mathematics (solving problems without calculation devices); and 4) self-checking-verifying of answers.

³ Coders were Ph.D. graduate students with extensive mathematics background in computation theory, undergraduate mathematical training and graduate training up through advanced econometric and psychometric statistical analyses.

⁴ Textbooks and their intended audiences tended to be different prior to 1930, often mixing language and numeracy skills for example, and it was not until the development of mass schooling in the U.S. by about 1930 that textbook designers explicitly focused on schooling all types of students in mathematics.

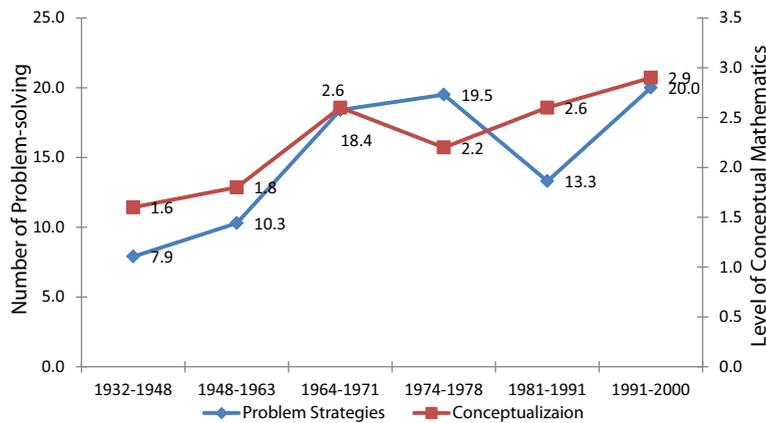


Fig. 5. Average number of different kinds of problem-strategies and average level of abstraction for basic arithmetic in 1st and 4th grade textbooks combined from 1932 to 2000.

2.3.2. Results

As shown in Fig. 5, the mean number of unique problem-solving strategies for arithmetic in first and four grade textbooks grew from about 8 to 20 during this time period, with a rapid increase to a mean of 18 strategies by the early 1970s.⁵ Similarly, the mean level of conceptual reasoning required to effectively use problem-solving strategies in these textbooks increased from little to none in the 1930s to moderate levels by the 1960s. Also as shown in Fig. 6, the frequency of one specific arithmetic problem strategy (i.e., reasoning) rose significantly from the 1930s to the 1960s, flattened out, and then substantially increased in the 1980s. Among the other high-CEF-demand strategies, the frequency of use of estimation drifted up from 1930 and then significantly increased from the early 1970s. In addition, after a dip from 1930 to the early 1970s, the use of mental mathematics and self-checking also significantly increased until the end of the century. Consequently, by the year 2000, young students were taught many different strategies to solve arithmetic problems, many of which demanded higher use of CEFs through greater incorporation of reasoning and other effortful cognitive skills than was the case in previous decades.

Furthermore, in widely used textbooks in the middle of the century, primary school students were presented with just three different strategies for solving multiplication problems: repeated addition of the same number, using related number facts to facilitate computation, and using arrays. By the 1980s, the same-grade students were taught five different strategies: using common multiples of numbers to facilitate computation; using zero place holders when multiplying by 10 or 100; estimating products; using place-value figures for multi-digit multiplication; and using a method for guessing, checking, and adjusting quotients to solve a problem. By 1999, a single textbook page often presented six or more strategies for solving arithmetic problems. Many of these additional strategies demanded more complex reasoning about mathematics; students were asked various meta-calculation questions such as why a particular strategy can derive the correct answer, how

mental math would get them to an answer, and how to do a self-check of one strategy versus another.

3. Discussion

Overall, the neurological, cognitive, and textbook-analysis results support the three-part schooling-cognition enhancement hypothesis, and for each set of findings there is supplemental corroborating evidence from related research.

The findings from Study 1 provide neurological evidence for H1 as brain activations associated with mathematical problem solving shared a common neural substrate with relational reasoning and other CEF capacities during the developmental years of schooling. These results are consistent with the notion that similar neurobiological mechanisms can serve the cognitive processes of mathematics and CEFs; and incremental mastery of one may well influence the other (e.g., Dehaene, Piazza, Pinel, & Cohen, 2003; Naghavi & Nyberg, 2005; Olesen, Westerberg, & Klingberg, 2004). Hence, math instruction throughout education may serve as a regular source of frontoparietal network stimulation that promotes focal parietal and frontal network activation and plasticity with increasing age and educational experience (Eslinger et al., 2009). Lastly, corroborating evidence from other MRI studies comparing unschooled and schooled individuals confirms that different brain activation patterns are recruited during language-mediated tasks lasting into adulthood (Carreiras et al., 2009; Castro-Caldas et al., 1999; Castro-Caldas, Petersson, Reis, Stone-Elander, & Ingvar, 1998; Petersson, Reis, Askelöf, Castro-Caldas, & Ingvar, 2000).

Evidence from Study 1 and others like it should not however be interpreted as an argument for *tabula rasa* and radical environmentalism in neuro-development. The neurobiological plasticity underlying human cognitive development is robust enough to thrive despite many biological and environmental hazards, but at the same time, there is increasing evidence that such development is invariably tethered to one's immediate environment and shapes the brain in specific ways, and multi-years of long hours in school is likely one such way (e.g., Eslinger et al., 2004). Yet, upward limits on the Flynn Effect likely exist, and the exact degree to which CEFs are

⁵ The dip in mean strategies to 13 in the 1980s was due to a "back-to-basics" movement that ended in the 1990s when the mean moved up to 20.

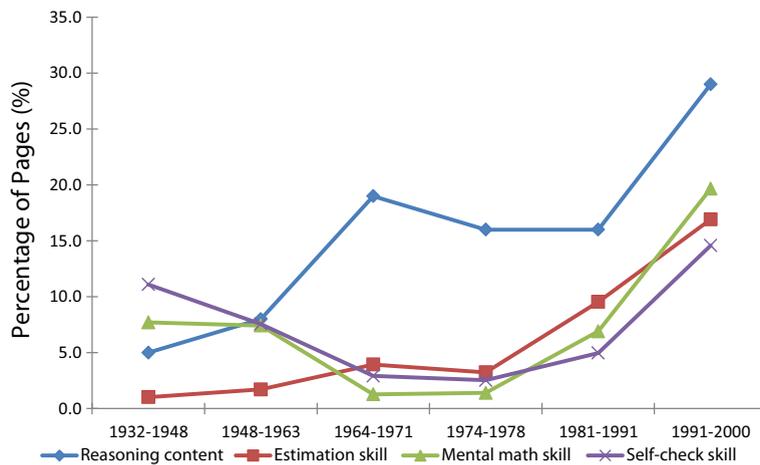


Fig. 6. The percentage of pages with reasoning, estimation, mental math, and self-check in the 1st and 4th grade textbooks combined, 1932–2000.

directly trainable is under debate (Owen et al., 2010; Teasdale & Own, 1987).

Although Study 2 could not randomly assign exposure to schooling, its design approximates an experiment on the cognitive effect of schooling. A significant improvement on earlier schooling-effect research, this study conditions on relevant covariates, including a concurrent measure of crystallized IQ. It also uses best scientific instrumentation of CEFs and is not confounded by significant neuro-maturation differences present in young students. The results show an influence of schooling on CEF enhancement; a one standard deviation increase in years of schooling results in a .63 standard deviation increase in CEFs. This schooling effect on CEF enhancement supports the evidence from prior comparisons of schooled and unschooled that have used more restricted designs (Burrage et al., 2008; Christian, Bachman, & Morrison, 2001; Cole, 1996; Stevenson, Chen, & Booth, 1990). In addition, there is evidence that certain cognitive competencies of U.S. students (ages 9, 13, 17) have increased from 1971 to 2008; similarly, findings of a recent Flynn Effect in South Korea are consistent with that country's late development of mass education (Rindermann & Thompson, 2013; te Nijenhuis et al., 2012).⁶ Without random assignment to school exposure, however, there is always the possibility of a selection effect and sources of endogeneity contaminating the results. But in this case, since all Study-2 participants reported starting and stopping their schooling because of the vicissitudes of subsistence-level farming, finances, access, or gender preferences instead of pre-school cognitive endowments, there is no evidence that variation in attendance and duration was confounded by a cognitive-selection effect or an unmeasured cognitive factor. And these findings are similar to those reported by Peters et al. (2010) in a similar study. In addition, Richards, Power, and Sacker (2009)

conclude from a series of analyses of British adults that education is likely a causal influence on later-life cognition even after conditioning on social class of origin, childhood intelligence, and subsequent occupational attainment. It should be noted that H2 does not imply that schooling is the sole cause of CEF enhancement as genetic and non-schooling environmental factors can also influence neuro- and cognitive-development. Education also may influence cognition through other mediating pathways, such as the possible influence of rising educational attainment of mothers on recently reported secular growth in pre-schoolers' IQ (Lynn, 2009b). Lastly, some inconsistent trends between adults' and children's gains point to a need for more research on occupational effects and multiplier effects between heritability and access to environments such as high quality schooling (Flynn, 2012).

As noted above, one reason that the education revolution was prematurely rejected as a plausible cause of rising population CEFs is the presumption that either expanding education has produced less cognitively demanding schooling or that the discipline-based content of schooling does not change much from generation to generation. For primary school mathematics, the results of Study 3 indicated that the overall cognitive demand of textbook material increased significantly. Starting in the mid-1960s, and increasing through the remainder of the century, young students were taught with material necessitating more effortful reasoning approaches to arithmetic requiring high-CEF-demand strategies. Along with the growing cognitive demand of material, another study finds that moving down in grade each decade since 1960 textbooks introduce significantly more mathematical sub-topics with less required rote drill at (Baker et al., 2010). Additionally, a recent parallel historical study of reading textbooks reports evidence of rising cognitive demand in what young students are asked to do with a text (Stevens, in press). Also, a similar shift was found in a comparison of the cognitive skills tested in high-stakes school examinations early and late in the 20th century, and changing goals of language instruction (Cha, 1991; Genovese, 2002). An important caveat to these findings that requires future research is evidence indicating that, despite increased instruction, mathematics reasoning skill enhancement among American students declines late in secondary school (Flynn, 2012).

⁶ There was evidence that the verbal ability of American secondary school students taking the SAT and ACT decreased in the 1960s and 1970s corroborated by analysis from the GSS, but subsequent analysis finds: 1) that this trend reversed by 1980s; 2) the trend is in part a function of changes in the population of students taking these tests; and 3) NAEP scores from national probability samples of students find increasing verbal and mathematical abilities since 1971 (Powell & Steelman, 1996; Rindermann & Thompson, 2013).

4. Implications of the schooling-cognitive enhancement hypothesis

If future research supports the current findings that the education revolution is one cause of rising population IQ over the past century, there are several major implications. Emerging conclusions from the study of human genetics and the ontogenesis of neuro-cognition indicate that neurological and cognitive development occurs through activity-dependent neural plasticity (e.g., Quartz & Sejnowski, 1997). In other words, inheritance of genetic potential for intelligence and relevant environmental factors interact to produce the complex phenotype of mental behavior and increasingly it makes less sense to think of genes and environments as independent causes. Additionally, this process is not only triggered by early nurturance since it is known to be sensitive to social environments throughout the life course (Baker et al., 2012). Therefore, schooling is not just an arena in which mental skills are employed, it can be ontogenetically fundamental to the creation of the neuro-architecture required for the expression of these skills within genetic boundaries. At the same time, the world prior to mass schooling clearly did not operate on extremely low levels of intelligence nor will mostly the brilliant populate the future. What the education revolution has done instead is to direct large amounts of children's and youth's time and motivation on types of cognitive skills. A laser-like focus on cognitive activities along with the growing social meaning of school achievement creates a culture of what can be called "academic intelligence" that pervasively pushes aside and devalues many other kinds of human capabilities (Baker, 2014).

What these findings might mean for understanding the Flynn Effect depends on contrasting scientific positions on the causes of human intelligence in individuals and within populations over time. From the activity-dependent neural plasticity perspective, the findings nominate formal education as a major environmental impact on neuroplasticity of CEFs that within a genetic range can produce historical change in the level of a phenotypic expression within populations. From the traditional perspective on g that is more skeptical of the meaning of the Flynn Effect to begin with the findings can be interpreted several ways. First, if the Flynn Effect does not reflect a substantial historical change in the mean level of g , then formal education should be added to the list of causes of a trivial trend such as measurement artifacts, changing test-taking strategies, and item exposure (e.g., Rodgers, 1998; Wicherts et al., 2004). The fMRI findings in particular, however, point to a fundamental neurological response to repeated exposure to academic activities that challenge such an interpretation. Second, if CEFs are relatively narrow abilities that while favorable to environmental effects are not the foundation for g , then the results are not very relevant to understanding intelligence (e.g. Armstrong & Woodley, 2014). Recent research claims that mean simple reaction time (RT) has slowed over a long historical period and thus reflects a dysgenic fertility trend in inherited intelligence in the population that has been masked by the more environmentally sensitive Flynn Effect on IQ scores (Woodley, te Nijenhuis, & Murphy, 2013). However, new counterfactual evidence challenges this claim, and CEFs are known to generalize to performance on a wide range of complex everyday tasks and are unlikely narrow abilities

(Gottfredson, 1997; Nettelbeck, 2014; Nettelbeck & Wilson, 2004). Lastly the findings beg the question: is it possible to separate empirically genetic capacity for g and its phenotypic expressions? The exact relationship between CEFs and overall intelligence is a central topic for future research, just as the relationship among the Flynn Effect, g , and dysgenic outcomes continues to be debated (e.g.; Lynn, 2011; te Nijenhuis & van der Flier, 2013). But at the very least, the results indicate that the shared meanings, reinforcements, and punishments of social institutions can interact with genetic endowments to enhance capabilities of individuals, and thus population patterns can correspond with the historical waxing and waning of social institutions.

One interesting example of this awaiting further research is that as mass schooling legitimates a link between cognition and social outcomes, cognitive skills may vary less across demographic backgrounds and occupations among cohorts born after the intensification of the American version of the education revolution compared to those born before (Weakliem, McQuillian, & Schauer, 1995). Furthermore, it is increasingly clear that people's cognitive skills interact with everyday life in ways that influence important outcomes, and the Flynn Effect may cause widespread consequences as a result (Howard, 1999). Recent epidemiological research finds a positive association between intelligence (including CEFs) and health outcomes (e.g., Rindermann & Gerhard, 2009). If CEFs are profoundly shaped and enhanced by education, these skills may be a plausible mediator of the widely reported associations between education and individuals' health, fertility, and longevity. It also opens a distinctly cognitive perspective on education's observed association with demographic processes and health outcomes at both the individual and population levels (e.g., Baker et al., 2011; Dieckmann et al., in press; Lutz & Samir, 2011; Peters et al., 2010; Rindermann & Gerhard, 2009).

The debate over the *Bell Curve* hypothesis is equally transformed by the results here. If mass schooling is one major cause of the Flynn Effect, then the basic premise of the *Bell Curve* is challenged. Herrnstein and Murray argued that social inequality is heavily related to intelligence assumed to be completely genetically determined (1994). Therefore, inequality is not amenable to governmental attempts to lessen it through improving the social environment for the more disadvantaged (Fischer et al., 1997). But with new research, this argument has proved to be incorrect genetically, neurologically, and sociologically. First, although it is true that individual variation in basic capacity for intelligence is genetically inherited to a considerable degree, it is not fully inheritable. Second and as suggested here, intelligence as a phenotype is formed through an interaction between genetic endowment and environment, and the present results indicate that education, essentially a massive social program, can be a fundamental factor in the environment's influence. Further, there is evidence that the relative impact of heritability and environment on intelligence can vary with parental SES, with more impact of environment than heritability among more disadvantaged families (Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003).

Formal education is certainly not the only social institution with the potential to shape cognitive development of individuals and changes across cohorts. Early parenting, social class habitus, informal learning environments and so forth

are important to consider in future research. And, historical inequality across social groups in access to, and quality of, schooling should be considered in future research on the Flynn Effect. But the expansion and intensification of formal education over the 20th century is likely responsible, at least in part, for the rise of CEF skills among successive cohorts.

Acknowledgments

The authors acknowledge funding from the National Science Foundation (SES-0826712; SES-1155924).

Appendix A. fMRI procedures

A.1. Preparation and positioning

fMRI studies were conducted on a 3.0 T MRI scanner. Stimuli were presented to participants through VisuaStim Digital Glasses (Resonance Technology Inc., Northridge, CA) and responses were recorded through a handheld device. All subjects were first introduced to the tasks and response device in out-of-magnet training that included introduction and instruction slides as well as sample trials of baseline and experimental stimuli on a laptop computer. Questions about the tasks and procedures were clarified with subjects until they fully understood the tasks for the fMRI session. Training was geared at alleviating any anxiety during fMRI study and thoroughly familiarizing participants with the modes of stimulus presentation, task requirements, and response options. Instruction slides were repeated in-magnet and task readiness cues were presented through the 2-way intercom.

Participants lay supine in a head restrainer that minimized motion and provided precise positioning and comfort. A boxcar fMRI paradigm was used, which consisted of interleaved time intervals of baseline and cognitive activation. During scanning, participants were instructed to respond to visual stimulation by pressing either the left or right button with their respective thumb on a 2-button handheld device.

A.2. Image acquisition

The Philips Intera 3.0 T whole-body scanner with SENSE parallel imaging ability (Philips Medical Systems, the Netherlands) was used for fMRI data acquisition. A T_1 -weighted three-dimensional image (3-D DEFT) was first acquired for anatomical structure with time of repetition (TR)/echo time (TE)/flip angle (FA) = 25 ms/4.6 ms/30°, FOV = 230 × 230 × 140 mm³, acquisition matrix = 240 × 180 × 140, and SENSE factor = 2. Echo planar imaging (EPI) that was used for fMRI data acquisition with TR = 3000 ms, TE = 35 ms, FA = 90°, 25 axial slices, slice thickness = 5 mm with no gap in between, FOV = 230 × 230 mm², acquisition matrix = 80 × 80, SENSE factor = 2. Two study paradigms were administered in each participant session. For the relational reasoning paradigm, 166 images were acquired during the execution of 3 cycles of stimulation and baseline.

A.3. Data analysis

fMRI image data were processed with SPM2 software (Wellcome Department of Cognitive Neurology, London, UK)

implemented in Matlab (Mathworks, Inc.). The first 4 images of each fMRI data set were discarded to remove the initial transit signal fluctuations and subsequent images were re-aligned within the session to remove any minor movements. No movement >2 mm in any direction of translation or 2° in any direction of rotation was observed. The T_1 -weighted high-resolution anatomical images were co-registered with fMRI images and spatially normalized according to the Montreal Neurological Institute brain template. The time-course images were normalized using the same normalization parameters and then smoothed with a $5 \times 5 \times 12.5$ mm³ (full width at half maximum) Gaussian smoothing kernel. A statistic parametric map (SPM) was generated for each subject under each condition by fitting the stimulation paradigm to the functional data, convolved with a hemodynamic response function. The pixels representing the active regions were overlaid on the 3-D T_1 -weighted anatomic image in Talairach coordinates. Brain activations generated by the relational reasoning and calculation (experimental) tasks were contrasted with their respective baseline conditions of perceptual matching and simple coin counting, isolating cognitive processes of relational problem solving and calculation.

Conjunction analyses were undertaken in order to identify those regions of activation that were shared by both tasks. Results were explored with a one sample *t*-test, with *p* values varying from .01 to .05.

References

- Arдила, A., Bertolucci, P.H., Braga, L.W., Castro-Caldas, A., Judd, T., Kosmidis, M.H., et al. (2010). Illiteracy: The neuropsychology of cognition without reading. *Archives of Clinical Neuropsychology*, 25, 689–712.
- Armstrong, E.L., & Woodley, M.A. (2014). The rule-dependence model explains the commonalities between the Flynn effect and IQ gains via retesting. *Learning and Individual Differences*, 29, 41–49.
- Baker, D. (2014). *The schooled society: The educational transformation of global culture*. Palo Alto, CA: Stanford University Press.
- Baker, D., Nripe, H., Cummings, E., Collins, J., Leon, J., Blair, C., et al. (2010). One hundred years of American primary school mathematics: A content analysis and cognitive assessment of textbooks from 1900 to 2000. *Journal for Research in Mathematics Education*, 41, 383–423.
- Baker, D.P., Leon, J., Smith Greenaway, E.G., Collins, J., & Movit, M. (2011). The education effect on population health: A reassessment. *Population and Development Review*, 37, 307–332.
- Baker, D., Salinas, D., & Eslinger, P. (2012). An envisioned bridge: Schooling as a neurocognitive developmental institution. *Developmental Cognitive Neuroscience*, 25, S6–S17.
- Bills, D. (2004). *Sociology of education and work*. Malden, MA: Blackwell.
- Blair, C., Gamson, D., Thorne, S., & Baker, D. (2005). Rising mean IQ: Cognitive demand of mathematics education for young children, population exposure to formal schooling, and the neurobiology of the prefrontal cortex. *Intelligence*, 33, 93–106.
- Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, 19, 273–293.
- Burrage, M.S., Ponitz, C.C., McCreedy, E.A., Shah, P., Sims, B.C., Jewkes, A.M., et al. (2008). Age- and schooling-related effects on executive functions in young children: A natural experiment. *Child Neuropsychology*, 14, 510–524.
- Carreiras, M., Seghier, M.L., Baquero, S., Estévez, A., Lorano, A., Devlin, J.T., et al. (2009). An anatomical signature for literacy. *Nature*, 461, 983–986.
- Castro-Caldas, A., Miranda, P.C., Carmo, I., Reis, A., Leote, F., & Ribeiro, C. (1999). Influence of learning to read and write on the morphology of the corpus callosum. *European Journal of Neurology*, 6, 23–28.
- Castro-Caldas, A., Petersson, K.M., Reis, A., Stone-Elander, S.M., & Ingvar, M. (1998). The illiterate brain: Learning to read and write during childhood influences the functional organization of the adult brain. *Brain*, 121, 1053–1063.
- Ceci, S. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, 27, 703–722.

- Cha, Y. (1991). Effect of the Global System on Language Instruction, 1850–1986. *Sociology of Education*, 64(1), 19–32.
- Christian, K., Bachman, H.J., & Morrison, F.J. (2001). Schooling and cognitive development. In R.J. Sternberg, & R.L. Grigorenko (Eds.), *Environmental effects on cognitive abilities* (pp. 287–335). Mahwah, NJ: Erlbaum.
- Cliffordson, C., & Gustafsson, J. (2008). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence*, 36, 143–152.
- Clouston, S., Kuh, D., Herd, P., Elliott, J., Richards, M., & Hofer, S.M. (2012). Benefits of educational attainment on adult fluid cognition: International evidence from three birth cohorts. *International Journal of Epidemiology*, 41, 1729–1736.
- Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Belknap Press.
- Deco, G., Rolls, E.T., & Horowitz, B. (2004). “What” and “where” in visual working memory: A computational neurodynamical perspective for integrating fMRI and single-neuron data. *Journal of Cognitive Neuroscience*, 16, 683–701.
- Dehaene, S. (1996). The organization of brain activations in number comparison: Event-related potentials and the additive-factors method. *Journal of Cognitive Neuroscience*, 8, 47–68.
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20, 487–506.
- Delis, D.C., Kaplan, E., & Kramer, J.H. (2001). *Delis–Kaplan Executive Function System (D-KEFS) technical manual*. San Antonio, TX: The Psychological Corporation.
- Dieckmann, N.F., Peters, E., Leon, J., Benavides, M., Baker, D.P., & Norris, A. (2015). The role of objective numeracy and fluid intelligence in sex-related protective behaviors. *Current HIV Research* (in press).
- Downey, D., von Hippel, P., & Broh, B. (2004). Are schools the great equalizer? Cognitive inequality during the summer months and the school year. *American Sociological Review*, 69, 613–635.
- Eslinger, P., Blair, C., Wang, J., Lipovsky, B., Realmuto, J., & Baker, D. (2009). Developmental shifts in fMRI activations during visuospatial relational reasoning. *Brain and Cognition*, 69, 1–10.
- Eslinger, P.J., Flaherty–Craig, C.V., & Benton, A.L. (2004). Developmental outcomes after early prefrontal cortex damage. *Brain and Cognition*, 55, 84–103.
- Fischer, C.S., Hout, M., Jankowski, M.S., Lucas, S.E., Swidler, A., & Voss, K. (1997). *Inequality by design: Cracking the bell curve myth*. New York: Academy of Political.
- Flynn, J.R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95, 29.
- Flynn, J.R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171.
- Flynn, J.R. (2009). *What is intelligence?* New York: Cambridge University Press.
- Flynn, J.R. (2012). *Are we getting smarter? Rising IQ in the twenty-first century*. UK: Cambridge University Press.
- Genovese, J. (2002). Cognitive skills valued by educators: Historical content analysis of testing in Ohio. *The Journal of Educational Research*, 96, 101–114.
- Gottfredson, L.S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24, 79–132.
- Hatch, S.L., Feinstein, L., Link, B.G., Wadsworth, M.E., & Richards, M. (2007). The continuing benefits of education: Adult education and midlife cognitive ability in the British 1946 birth cohort. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 62, 404–414.
- Hayes, T., Petrov, A., & Sederberg, P. (2015). Do we really become smarter when our fluid-intelligence test scores improve? *Intelligence*, 48, 1–14.
- Herstein, R., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.
- Horgan, J. (1995). Get smart, take a test. *Scientific American*, 273, 12–14.
- Howard, R.W. (1999). Preliminary real-world evidence that average human intelligence really is rising. *Intelligence*, 27, 235–250.
- Jensen, A.R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Kaufmann, L., Wood, G., Rubinstein, O., & Henik, A. (2011). Meta-analyses of developmental fMRI studies investigating typical and atypical trajectories of number processing and calculation. *Developmental Neuropsychology*, 36, 763–787.
- Kline, R.B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Kroger, J.K., Sabb, F.W., Fales, C.L., Bookheimer, S.Y., Cohen, M.S., & Holyoak, K.J. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex*, 12, 477–485.
- Luria, A.R. (1976). *Cognitive development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.
- Lutz, W., & Samir, K.C. (2011). Global human capital: Integrating education and population. *Science*, 333(6042), 587–592.
- Lynn, R. (1989). A nutrition theory of the secular increases in intelligence; positive correlations between height, head size and IQ. *British Journal of Educational Psychology*, 59, 372–377.
- Lynn, R. (2009a). Fluid intelligence but not vocabulary has increased in Britain. *Intelligence*, 37, 249–255.
- Lynn, R. (2009b). What has caused the Flynn effect? Secular increases in the development quotients of infants. *Intelligence*, 37, 249–255.
- Lynn, R. (2011). *Dysgenics: Genetic deterioration in modern populations*. London: Ulster Institute for Social Research.
- Lynn, R. (2013). Who discovered the Flynn effect? A review of early studies of the secular increase of intelligence. *Intelligence*, 41, 765–769.
- Lynn, R., & Meisenberg, G. (2010). National IQs calculated and validated for 108 nations. *Intelligence*, 38, 353–360.
- Martinez, M. (2000). *Education as the cultivation of intelligence*. Mahwah, NJ: Erlbaum.
- Martorell, R. (1998). Nutrition and the worldwide rise in IQ scores. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 183–206). Washington, DC: American Psychological Association.
- McGrew, K.S., Wood, R.W., & Mather, N. (2001). *Woodcock Johnson III*. Itasca, IL: Riverside.
- Naghavi, H.R., & Nyberg, L. (2005). Common fronto-parietal activity in attention, memory, and consciousness: Shared demands on integration? *Consciousness and Cognition*, 14, 390–425.
- Neisser, U. (1998). Rising test scores and what they mean. In U. Neisser (Ed.), *The rising curve: Long term gains in IQ and related measures*. Washington, DC: American Psychological Association (PAGE NOS. NEEDED).
- Nettelbeck, T. (2014). Smarter but slower? A comment on Woodley, te Nijenhuis & Murphy (2013). *Intelligence*, 42, 1–4.
- Nettelbeck, T., & Wilson, C. (2004). The Flynn effect: Smarter not faster. *Intelligence*, 32, 85–93.
- Nisbett, R.E. (2009). *Intelligence and how to get it*. New York: Norton.
- Nisbett, R.E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D.F., et al. (2012). *Intelligence: New findings and theoretical developments*. *American Psychologist*, 67, 130–159.
- Okagaki, L., & Frensch, P.A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of Applied Developmental Psychology*, 15, 33–58.
- Olesen, P.J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, 7, 75–79.
- Owen, A.M., Hampshire, A., Grahn, J.A., Stenton, R., Dajani, S., Burns, A.S., et al. (2010). Putting brain training to the test. *Nature*, 465, 775–778.
- Peters, E., Baker, D.P., Dieckmann, N.F., Leon, J., & Collins, J. (2010). Explaining the effect of education on health: A field study in Ghana. *Psychological Science*, 21, 1369–1376.
- Petersson, K.M., Reis, A., Askelöf, S., Castro-Caldas, A., & Ingvar, M. (2000). Language processing modulated by literacy: A network analysis of verbal repetition in literate and illiterate subjects. *Journal of Cognitive Neuroscience*, 12, 364–382.
- Powell, B., & Steelman, L. (1996). Bewitched, bothered, and bewildering: The use and misuse of state SAT and ACT scores. *Harvard Educational Review*, 66, 27–60.
- Price, C.J., & Friston, K.J. (1997). Cognitive conjunction: A new approach to brain activation experiments. *NeuroImage*, 5, 261–270.
- Quartz, S.R., & Sejnowski, T.J. (1997). The neural basis of cognitive development: A constructivist manifesto. *Behavioral and Brain Sciences*, 20, 556–596.
- Raven, J., Raven, J.C., & Court, J.H. (2000). *Standard progressive matrices*. San Antonio, TX: Harcourt.
- Richards, M., Power, C., & Sacker, A. (2009). Paths to literacy and numeracy problems: Evidence from two British birth cohorts. *Journal of Epidemiology and Community Health*, 63, 239–244.
- Rindermann, H., & Gerhard, M. (2009). Relevance of education and intelligence at the national level for health: The case of HIV and AIDS. *Intelligence*, 37, 383–395.
- Rindermann, H., & Thompson, J. (2013). Ability rise in NAEF and narrowing ethnic gaps? *Intelligence*, 41, 821–831.
- Rivera, S.M., Reiss, A.L., Eckert, M.A., & Menon, V. (2005). Developmental changes in mental arithmetic: Evidence for increased functional specialization in the left inferior parietal cortex. *Cerebral Cortex*, 15, 1779–1790.
- Rodgers, J.L. (1998). A critique of the Flynn effect: Massive IQ gains, methodological artifacts or both? *Intelligence*, 26, 337–356.
- Runquist, E.A. (1936). Intelligence test scores and school marks in 1928 and 1933. *School & Society*, 43, 301–304.
- Schaub, M. (2010). Parenting for cognitive development from 1950 to 2000: The institutionalization of mass education and the social construction of parenting in the United States. *Sociology of Education*, 83, 46–66.
- Spearman, C. (1923). *The nature of intelligence and the principles of cognition*. London: Macmillan.
- Stevens, R., Lu, X., Baker, D.P., Ray, M., Eckert, S., & Gamson, D. (2015). Assessing the cognitive demands of a century of reading curricula: An analysis of

- reading text and comprehension tasks from 1910 to 2000. *American Education Research Journal* (in press).
- Stevenson, H., Chen, C., & Booth, J. (1990). Influences of schooling and urban-rural residence on gender differences in cognitive abilities and academic achievement. *Sex Roles*, 23, 535–551.
- te Nijenhuis, Jan, & van der Flier, Henk (2013). Is the Flynn effect on *g*? A meta-analysis. *Intelligence*, 41(6), 802–807.
- te Nijenhuis, J., Cho, S. H., Murphy, R., & Lee, K. H. (2012). The Flynn effect in Korea: Large gains. *Personality and Individual Differences*, 53(2), 147–151.
- Teasdale, T.W., & Own, D.R. (1987). National secular trends in intelligence and education: A twenty-year cross-sectional study. *Nature*, 325, 119–121.
- Torche, F. (2011). Is a college degree still the great equalizer? Intergenerational mobility across levels of schooling in the United States. *American Journal of Sociology*, 117, 763–807.
- Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I.I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, 14, 623–628.
- U.S. Department of Education. National Center for Education Statistics (2006). The condition of education 2006. Retrieved from nces.ed.gov/pubs2006/2006071.pdf
- U.S. Department of Education. National Center for Education Statistics (2008). Digest of education statistics 2007. Retrieved from nces.ed.gov/pubs2008/2008022.pdf
- UNESCO (2002). Education for all: Is the world on track? Retrieved September 17, 2014, from unesdoc.unesco.org/images/0012/001290/129053e.pdf
- Van de Vijver, F.J., & Brouwers, S.A. (2009). Schooling and basic aspects of intelligence: A natural quasi-experiment in Malawi. *Journal of Applied Developmental Psychology*, 30, 67–74.
- Weakliem, D., McQuillian, J., & Schauer, T. (1995). Toward meritocracy? Changing social-class differences in intellectual ability. *Sociology of Education*, 68, 271–286.
- Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale: Revised*. San Antonio, TX: The Psychological Corporation.
- Wicherts, J., Dolan, C., Hessen, D., Oosterveld, P., van Baal, C., Boomsma, D., et al. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence*, 32, 509–537.
- Wilkinson, G.S. (1993). *Wide Range Achievement Test 3: Administration manual*. Wilmington, DE: Jastak Associates.
- Williams, W.M. (1998). Are we raising smarter children today? School- and home-related influences on IQ. In U. Neisser (Ed.), *The rising curve: Long term gains in IQ and related measures* (pp. 125–154). Washington, DC: American Psychological Association.
- Willis, S., Tennstedt, S., Marsiske, M., Ball, K., Elias, J., Koepke, K., et al. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *Journal of the American Medical Association*, 296, 2805–2814.
- Woodley, M. A., te Nijenhuis, J., & Murphy, R. (2013). Were the Victorians cleverer than us? The decline in general intelligence estimated from a meta-analysis of the slowing of simple reaction time". *Intelligence*, 41.6, 843–850.