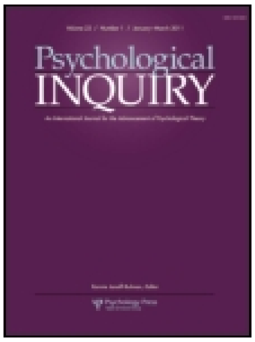

Routledge
Taylor & Francis Group

Psychological Inquiry

An International Journal for the Advancement of Psychological Theory

ISSN: 1047-840X (Print) 1532-7965 (Online) Journal homepage: http://www.tandfonline.com/loi/hpli20

No Population Is Frozen in Time: The Sociology of Intelligence

James R. Flynn

PSYCHOLOGICAL INQUIRY
2016, VOL. 27, NO. 3, 205–209
http://dx.doi.org/10.1080/1047840X.2016.1181518

No Population Is Frozen in Time: The Sociology of Intelligence

James R. Flynn

Department of Psychology, University of Otago, Dunedin, New Zealand

The target article by Kristof Kovacs and Andrew Conway (this issue) is of great importance. Of particular interest is its division of the common factor in IQ test performance into three elements: induction, working memory capacity, and executive functions. I divide my commentary as follows: the theory's superiority to Jensen's theory on the within-population (individual differences) area of intelligence, how the three elements listed illuminate what occurs in the between-population (sociological) area of intelligence, and how social trends over time generate novel hypotheses on both the individual differences and brain physiology areas.

Process Overlap Theory Versus *g*-Theory

Many deductions from Jensen's theory have now been falsified. All of them rest on the proposition that *g* is psychological process (with its own seat in the brain) tapped by all intelligence tests to obtain the positive manifold. I call this the "irreplaceable fuel" theory, which has four parts: (a) On the neural level, there is a mental energy (having to do perhaps with neural speed). (b) On the psychological level, this engenders a problem-solving capacity tapped by special problem-solving abilities (say those measured by different Wechsler subtests): (c) the degree to which it is tapped varies by subtest insofar as they pose problems of cognitive complexity and this is measured by their *g*-loadings, and (d) score gains over time that do not tally with a subtest hierarchy according to *g*-loadings are not true *g*-gains or intelligence gains.

They are gains of lesser consequence caused by social change (more familiarity with test content). They are not true intelligence gains because these come only with the upgrading of brains, thereby upgrading the quality of *g*. Improved brains are caused by factors that impinge on brains directly. For example, natural selection, hybrid vigor, better nutrition, either during pregnancy or in early childhood (breast-feeding), elimination of childhood diseases that damage brain development, and so forth. Needless to say, only a fragment of IQ gains over time can be explained by biological factors, and thus only this fragment is accorded the status of cognitive progress. For an analysis of Jensen's theory versus my own, see my recent book (Flynn, 2016, Part 2).[1]

I elaborate the criterion Jensen (1998) offered to determine whether score differences over time tallied with *g*. Take IQ gains from one generation to the next: You rank the 10 Wechsler subtests in order of the magnitude of the gains on each subtest, and then you rank the same subtests in order of the size of their *g*-loadings. The *g*-loading tell you the extent to which a particular subtest measured *g*, in the sense of what subtest was most predictive of the positive manifold, that is, the tendency of a good subtest performance to be sustained over all 10 subtests. Unless you find a robust positive correlation between the two hierarchies (biggest gain = highest *g* loading, etc.), the score gains do not constitute a *g* difference. IQ gains over time generally flunked this criterion and were therefore "hollow." Whatever fuelled them was not improved *g*-fuel.

We now know that whether IQ gains are *g* gains does not rob them of real-world significance. Coyle and Pillow (2008) showed that the cognitive skills measured by the Scholastic Aptitude Test predict university grades even after *g* has been removed. Woodley (2012a) showed that education in particular cultivates specialized patterns of cognitive abilities and that these improve independently of whether they correlate with *g*. Ritchie et al. (2015) were quite explicit: The association of education with improved cognitive performance is not mediated by *g*; education directly affects specific IQ subtests. Woodley (2012b) showed that the historical trend of IQ gains (which of course are not correlated with *g*) both parallels and predicts the growth in GDP per capita experienced by Western nations over the last 10 decades or so ($r = .930$). Meisenberg (2014) argued that over time we are accumulating "cognitive human capital" that is interdependent with economic growth.

There is an inference here I want to defend: Schooling promotes a variety of cognitive skills (*g* aside), and these promote economic progress. Note that the causal arrows could go in the opposite direction: x causes us to get richer, and we spend more on schools and get "smarter." My inference is more probable when we look at "lagged correlations" or what happens when the dimension of time is included. Ireland enhanced education, its tests scores rose, and its per capita gross domestic product rose above that of England—*in that order*. Finland enhanced education of its poorest students and duplicated Ireland's trend (Nisbet, in press).

Fox and Mitchum (2013) showed that IQ gains on Raven's reflect the kind of problems we can solve, despite the fact that they are not correlated with *g* and are not factor invariant. Using the Advanced Progressive Matrices test, Fox and Mitchum allow us to analyze what has altered in people's minds when one generation scores higher on Raven's than the last.

CONTACT James R. Flynn jim.flynn@otago.ac.nz Department of Psychology, University of Otago, PO Box 56, Dunedin 9054, New Zealand.

[1]I thank the syndics of the Cambridge University Press for permission to use material from James R. Flynn, *Does Your Family Make You Smarter? Nature, Nurture, and Human Autonomy*, Cambridge University Press, 2016.

The following analysis is in my language (reproduced from Flynn, 2012a, pp. 284–286). However, we met at the University of Richmond, and they confirmed that my interpretation is compatible with their analysis.

Some 115 years ago, people just beginning to enjoy modernity were still focused on the concrete objects of the real world. They wanted to manipulate the real world to their advantage, and therefore the representational images of objects was primary. If you are hunting you do not want to shoot a cow rather than a deer; if a bird is camouflaged in a bush, you flush it out so its shape can be clearly seen. Raven's poses a problem that is quite alien to your "habits of mind": You must divine relations that emerge only if you "take liberties" with the images presented. It is really a matter of perceiving analogies hidden behind distracters. I present a series of analogies (the first three are my own) to illustrate the point.

1. Dogs are to domestic cats as wolves are to (wild cats). Presented with these representational images people a century ago would have no difficulty.
2. ■ is to ♦ as ↑ is to (→) where the choices are ↑, →, ↖, and ↗. Here you must ignore everything about an image except its shape and position. Just as the square has been rotated a half turn, so has the arrow.
3. □ is to / as O is to (|) where the choices are ∅ - Θ -| - ⊗. Here you must ignore everything but the number of dimensions: The analogy compares two-dimensional shapes to one-dimensional shapes and all else is irrelevant. Representational images are of course three-dimensional, so such a contrast requires being well removed from them.
4. &#B is to B&# as T&T is to ##_ (enter what symbol fits). This is an item from Fox and Mitchum that illustrates the kind of analogical thinking you must do on the Advanced Raven's Progressive Matrices.

Note that the right answer in the fourth item has been left blank. Because no alternatives were presented to choose from, you had to deduce that "&" is the correct answer. I got it right, which was reassuring given that I was then 78 years old, by reasoning as follows. In the first half of the analogy, all that has altered is the sequence of symbols: labeling them 1, 2, 3, they have become 3, 1, 2. Applying that to the second half of the analogy, T&T changes to TT&. Clearly you are supposed to ignore the fact that the doubled letter (TT) has changed to a doubled symbol (##), so the right answer is ##&. This would really discriminate between the generations. We have moved far away from the "habit of mind" of taking pictorial images at face value; indeed, we are interested only in their sequence and treat images themselves as interchangeable if the logic of the sequence demands it.

The key is this: Anyone fixated on the literal appearance of the image "T," as a utilitarian mind would tend to be, would simply see no logical pattern. Contrast this with Wechsler Adult Intelligence Scale Vocabulary (here gains are large as distinct from Wechsler Intelligence Scale for Children Vocabulary). The etiology of enhanced scores over time would be quite different. People over time, thanks to the bonus of more education, simply accumulated a larger store of core vocabulary and got no bonus from the shift from utilitarian toward "scientific" thinking. Except of course for words that labeled abstractions (like species), which now appeared in the new subjects taught.

Fox and Mitchum (2013) classified Raven's items in ascending order of "relational abstraction," more specifically: "for analogical mapping when relations between objects are unrelated to objects themselves." Once again, in Example 4, the relationship can be derived only if one sees that a "T" does not have to retain its identity as a "T." Their core assumption was that "analogical mapping of *dissimilar* objects is more difficult than mapping *similar* objects" (italics mine). I certainly found this to be true. The fact "TT&" had to be translated into "##&" rendered the item harder to solve. And if I were my father (born in 1885), and wedded to taking images at face value for reasons of utility, I suspect I would have found it insuperable.

They analyze the performance of two samples of young adults tested in 1961 and circa 2006, respectively. They found that as the degree of deviation toward the abstract increased, certain items became less predictive of performance within the two generations than between the two generations.

We now know why Raven's scores are so sensitive to environmental change over time. Like our ancestors, we can still use logic to analyze the concrete world. But we have entered a whole new world that allows us to use logic on symbols far removed from the concrete world. We organize the concrete world using abstract concepts that are not represented there.

Premodern people see fish as having nothing in common with crows. You can eat one and not the other; one swims, the other flies. We use DNA analysis to divide living creatures into categories that are nonobservable but offer understanding, and this language has become that of every person who has been exposed to several years of formal schooling. We know that bacteria differ from one-celled animals, that whales are more akin to land animals than fish, and that the tiny hyrax is more akin to the huge elephant than to the rodents it resembles. We know that stars are different from planets (they look the same in the sky), and indeed, our whole picture of the universe (and even our approach to explaining human behavior) is based on logic and abstractions. We are exposed to the symbolism of algebra. No one has ever observed an "x."

In other words, using logic on symbols detached from concrete reality has become a habit of mind in no way alien to us. These skills are not merely useful in mathematics and science and computer programming (programmers do very well on Raven's). They help us to create (and comprehend) a nonrepresentational map of the London underground, or an organizational map that functionally relates the tasks a complex business organization performs. We are more ready to engage with Raven's because the rise of modernity altered our perspective. And the rise of modernity has occurred over only a few generations. Only a test that is sensitive to the new minds that modernity has put into our heads could measure something so malleable. Raven's, more than any other test, *is a barometer of the stages of modernity* and thus continues to play a crucial role in the study of intelligence.

Fox and Mitchum (2014) extended their analysis to Letter Series and Word Series and showed that the fact that the present generation has developed new habits of mind is the very reason gains are not factor invariant. Woodley, Figueredo, Ross, and Brown (2013) concluded that autonomous mental skills allow people to cognitively adapt to modernity and thus score higher on personality indexes. Flynn (2012a) showed that the fact that American adults with some tertiary education went from 12% to 52% between 1953 and 2007 registered as huge gains on the WAIS Vocabulary subtest. These were the equivalent of 17 IQ points (over 1 standard deviation). Irrespective of whether the overall pattern of American subtest gains correlated with *g*, this had real-world consequences: They could carry on different conversations and read a wider range of books. Flynn (2013) suggested how cognitive progress independent of *g* has enhanced moral maturity (but not political maturity).

There always was something odd about that the notion that performance gains must tally with complexity for the gains to have real-world consequences. Two basketball teams are evenly matched. The coach of one decides to drill his players on the fundamentals, layups, and foul shots, simple tasks that are less "basketball-*g*" loaded. Therefore, the performance gains they make do not correlate with a hierarchy of basketball-skill *g* loadings (no gains on complex tasks like fade-away jump shots). Yet there are real-world consequences: His team beats their rivals by 10 points.

Flynn, te Nijenhuis, and Metzen (2014) put a nail in *g*'s coffin. They compared the Wechsler subtests scores of typical subjects with those who suffered from iodine deficiency, prenatal cocaine exposure, fetal alcohol syndrome, and traumatic brain injury. The typical subjects were higher on every subtest. However, the magnitude of their advantages by subtest had zero correlation with the size of the subtest *g* loadings. It is difficult to deny that the typical subjects had a significant real-world cognitive advantage over the four comparison groups. This is not to say that their advantage was analogous to that of one generation over another. The latter was influenced by the new habits of mind that evolved over the 20th century.

Now process overlap theory puts a second nail in the coffin. It strips *g* of its central role in Jensen's theory of intelligence: that of a psychological process tapped by all IQ tests to obtain the positive manifold. It shows that *g* emerges *because* of the positive manifold rather than *explaining* it. The proffered explanation of the positive manifold involves three elements: induction, working memory capacity, and executive function. These overlap, and the combination is always to some degree involved in performing cognitive tasks. Better still, it adds specificity by identifying the central role played by *induction*. No one actually solves *g* problems (whatever they might be), that is, it is not a functional mental ability. But people do induction, and it is clear why Raven's is the best test of "*g*": It is a test of induction beyond all other tests (John C. Raven called it eduction).

There is nothing odd about why the three elements cohere. Working memory capacity is clearly a prerequisite of induction: The greater your capacity to hold abstract concepts in mind, the more you can look for relevant similarities and differences. Executive function in this context is the ability to exclude both cognitive and emotional interference with the inductive task at hand. It is clearly a prerequisite for both induction and high working memory—and indeed, the solution of any other cognitive task.

One thing troubles me: the Wechsler subtests scores of typical subjects and those who suffered from iodine deficiency, prenatal cocaine exposure, fetal alcohol syndrome, and traumatic brain injury. Although typical subjects were higher on every subtest, the magnitude of their advantages by subtest had zero correlation with the size of the subtest *g* loadings. If we substitute for *g* the three-factor concept of induction, working memory capacity, and executive function, should there not be a correlation between the extent to which this package is relevant to the subtest and the score difference between normal and damaged subjects? Unless these maladies collectively (and indeed virtually singly) damage the prefrontal lobes in a way that somehow cancels out their differential contribution to the cognitive task set by the different subtests, perhaps by reducing its contribution in all cases to a minimum. This does not seem very plausible, and the authors may wish to comment.

Sociology and the Three Elements

What goes on in people's minds as they solve cognitive problems is a product of the kind of person they are in a particular social setting. Kovacs and Conway (this issue) confine themselves to a within-generation analysis (the common factor weaker at high levels of ability) with only a nod at between-generation analysis (the common factor weakens as generations produce more people of high ability). There is one exception: They imagine the difference between a normal and a color-blind population when they try to solve the colored version of Raven's Progressive Matrices, with the former population largely defeated by the test.

Different societies and different stages of society on the path to modernity alter the hierarchy of problems that are considered important and the habits of mind of the people who try to solve them. They produce radically different populations not unlike the difference between those who are color-blind and those who are not. Moreover, going from one population to another affects the balance between the three elements of induction, working memory capacity, and executive function.

Contrary to Jensen, I make these assumptions: (a) The brain is like a muscle and is modified by exercise; (b) Societies (and generations) have very different hierarchies as to what problems are most important; (c) Practice at solving these problems create different "habits of mind" suited to solving problems in order of importance; and (d) These habits of mind alter how induction, working memory capacity, and executive functions interact. To elaborate, people in 1900 did not need to confront everyday problems that required these habits of mind: taking the hypothetical seriously, using abstract concepts to classify, using logic to analyze relation between such concepts. Therefore, when confronting the inductive tasks of Raven's they were like the color-blind confronting the colored matrices, except worse: Not only were their minds unprepared for the inductive tasks, but also they could not see the point of them, which would undermine their executive capacity to ignore distractors. As to whether they had lower working memory capacity than we do, who knows? I cannot estimate whether we need to hold more things in mind to analyze the relationship between abstractions than to analyze the relationships of coping with everyday life.

This is an example drawn from our own society as it progresses toward modernity, but other preindustrial societies also rank the importance of cognitive tasks differently than we do. Australian Aboriginal society put a high premium on "map reading," that is, noting signals of the presence of water and game on the horizon and calculating the distance that must be traversed. Thus, they would put map reading at the top of an importance of cognitive skills hierarchy and inductive analysis of abstractions would hardly count. I do not say there would be no input from induction—you use induction to some degree in everyday life—but I suspect the input would be limited. Even in our own society the balance between map reading and induction has probably altered over time. When people began to pilot autos, they got more practice in map reading. When only the rich could own cars, those only who had developed their inductive capacity by more formal education would drive, and this would inflate the correlation between induction and mapping. Then the poor got cars, which would lower it. Now cars have road-trip planners or an automatic guidance system that should put map reading problems much lower on our scale of priorities and return us to the pre-car state.

The best illustration of how executive functions correlate with induction arises from an analysis of a consequence of cognitive progress often not perceived, namely, its role in promoting moral progress. Remember that the modern mind broke its ties with the concrete world, the dominant theme as late as 1900, and asked us to take the hypothetical seriously and use logic to analyze abstract concepts. How did these habits of mind take moral reasoning away from the Stone Age of simply accepting the bias and cruelty of the past?

First, there is taking the hypothetical seriously. When combating racism, taking the hypothetical seriously is the foundation of mature moral argument. In 1955, when Martin Luther King began the Montgomery bus boycott, young men of my acquaintance, home from college at 21, had dialogues with their parents or grandparents. Question: "What if you woke up tomorrow and had turned Black?" Reply: "That is the dumbest thing you have ever said, who do you know that turned Black overnight?" My father believed that problems had to be grounded in the real world to take them seriously and had no room for hypothetical problems.

As for nationalism, my *Beyond Patriotism* (Flynn, 2012b) tries to diagnose the retreat from patriotism by some of the American public between World War II and today. Try this question: "What if your home was hit by a drone because someone nearby was sheltering a Taliban?" Or better: "If a war killed so many foreigners to save 3,000 Americans, where would you fall off the boat: at 10,000 or 100,000 or one million?" The answer tends to divide the youth from the aged (the latter: "Their government protects them and our government protects us"). Voltaire said that all man's reason flies before a drum. Well, it depends on how much reason and how loud the drum.

Today we use logic to analyze abstract concepts. This is a powerful weapon against local norms that incorporate the cruelty of the past as a residue. An Islamic father (guided by local norms, not the Koran) shocks the world when he kills a daughter because she has been raped. We would ask: "What if you had been knocked unconscious and sodomized?' He is unmoved. He sees moral maxims as concrete things, no more subject to logic than any other concrete thing like a stone or a tree. He does not see them as universals to be generalized by logic. Today the tendency is to express moral maxims as generalizations and try to make them logically consistent with one another. Question for one of my students: You say we should never judge the customs of another culture, yet you are also an advocate of women's rights. What do you say about the practice of female circumcision? Whatever the conclusion, this is a far cry from primitive moral reasoning.

In other words, the new habits of mind did not merely help us to adapt to modernity. They also taught us how to modify the modern world thanks to more mature moral reasoning. They taught us to stride toward freedom with Martin Luther King and take seriously the "collateral damage" of killing foreigners in Vietnam and Iraq and Afghanistan. No general today would talk about "bombing the Vietnamese back to the Stone Age."

This makes it seem as if the evolution of society toward modernity has made the use of induction on moral problems merely a matter of developing new habits of mind, ones that are friendlier to logical analysis. However, the social setting has a profound influence on the role of executive functions. An affluent resident of an area in the Middle East or Africa may have had formal education, and thus modern habits of mind, but also come from a family dominated by an inherited sexist morality. The stress placed on his executive functions to banish emotions irrelevant to the application of logic to a moral question may be extraordinary compare to our own: The raped girl just seems somehow tainted. The same is true of someone who comes from a family dominated by racial prejudice: Real-world Blacks just seem alien in a way that impedes analysis based on the traits of hypothetical Blacks.

Intelligence and the Three Areas

In the area of explaining intelligence gains over time, causal explanation involves several levels: (a) Ultimate causes are the industrial revolution and the resulting trend toward modernity; (b) Intermediate causes are the effects of industrialization on society, more education, emancipation of women, smaller families (with a better adult to child ratio), more cognitively demanding jobs, more cognitively demanding leisure, a new pictorial and symbolic world from television and the Internet, better nutrition, and medical advance; and (c) Proximate causes have to do with how people's minds altered, so that in the test room they could do better when taking IQ tests (e.g., new habits of mind).

The Dickens/Flynn (Dickens & Flynn, 2001) model predicts that the size of the IQ advantage between generations will vary depending on the age at which we compare a later cohort (say those born in 1936) with an earlier cohort (say those born in 1921). Both of these groups live their own lives. During those lives the causal factors that differentiate the later from the earlier cohort vary greatly. This means that the IQ gap that separates the two will vary in magnitude with age according to the potency of the differential factors that kick in at each age. This prediction remained only a prediction until a recent study. As Staff, Hogan, and Whalley (2014) say, their study is the first to compare two cohorts at two different ages.

The Lothian Birth Cohorts were born in 1921 and 1936, respectively. They included almost every child born in Scotland in those years (and still attending school there at the age of 11). Both were tested on Raven's Progressive Matrices: The later cohort outscored the earlier by 3.7 IQ points at age 11 and by 16.5 IQ points at the age of 77. The difference is huge: The rates of gain differ at 0.247 points and 1.100 points per year over a period of 15 years. If anything the gain in old age is an underestimate: The earlier cohort lost more people by death (earlier death is negatively correlated with IQ) than the later. The differing gains must reflect the relative potency of the causal factors that separated the cohorts at those two ages. What might these be?

When you test two cohorts at the age of 11, they both have approximately the same number of years of formal schooling and this serves as a leveler: The small IQ gap (and vocabulary gap) would reflect only the fact that the later cohort came from homes a bit higher in socioeconomic status and any progress made in the quality of schooling. The IQ gap doubles at the age of 21 and, indeed, the Vocabulary gap quadruples: This is thanks to more students going on to tertiary education; the later cohort would have more years of formal education. By age 35, the influence of more schooling would have faded in favor of the later cohort working at more cognitively demanding jobs. No data reveal whether this would confer a greater or lesser advantage than was present in the university years.

At the age of 70, one might anticipate a lessening of the gap, as both cohorts would have retired from work—*except* that the later cohort would be far more healthy and alert. Modern medicine has alleviated the many of the illnesses of old age, and older people today have a better diet and do more exercise (I still run at 82, and my father took no exercise after age 14). Elderly people also have leisure activities that are more cognitively demanding. At age 77, we have real data. We know that the three factors named produce a huge gap (16.50 points for two cohorts only 15 years apart), a gap unlikely to be matched at any earlier age.

I have often rejected the hypothesis that generational IQ gains reflect gains in health and nutrition, at least in advanced nations since 1950. This was because we were looking for them in the wrong place: We thought they would weigh in at the beginning of life (they do not); rather they weigh in at the end. At any rate, we now know that Raven's is not merely sensitive to the global environment enriched by modernity. It is also sensitive to each and every one of the particular factors that have triggered IQ gains over time. This has implications not only on between-generation IQ differences but also on individual differences within a cohort. If one person gets more formal schooling than another, or a more demanding job than another, or better diet and medical care in old age, they will at the appropriate stage of life have an IQ advantage.

By their very nature, theories of brain physiology would ideally accommodate both individual differences and the evolution of cognitive abilities over time. First, we want to map the areas/networks that are activated when people perform various cognitive skills; then we will want to observe differences in those areas/networks that rank people's performance for each cognitive skill. In principle, brain physiology should also illuminate cognitive trends from one generation to another. It is a plausible hypothesis that as people began to drive motorcars, more mapping exercise enlarged the hippocampus between 1900 and today, and that the introduction of automatic guidance systems will erode the size of the hippocampus in the future. We must wait for data about the future but could project back into the past by studying drivers versus nondrivers—or ethnic groups that do not drive cars (the Amish).

Integration of All Areas

We want an adequate theory of intelligence in the area of brain physiology. However much we may succeed, we will have to resist the temptation of reductionism. Physiology cannot replace psychology and sociology in the sense that we will still need causal explanations in all three areas of the study of human intelligence. Physiology may be able to predict exactly who will be the best basketball player, but we still need to know why someone is doing something as trivial as running around a court to try to throw a ball through a hoop, and why basketball became more popular after World War II, so that greater participation rose and triggered a huge rise in standards of performance.

Kovacs and Conway's primary contribution is in the area of individual differences. At times, they say that the emergence of a positive manifold is a function of the psychological processes of individuals solving problems. I see no reason to assume that this implies that scholars can neglect the fact that individuals are the product of different social circumstances, and that this affects how they solve problems. Which is to say it does not assume we can neglect the sociology of intelligence. Nonetheless, I want to emphasize that a comprehensive understand of intelligence must integrate all three areas. What we think we know about individual differences will always be qualified by what is true about both the brain and society.

References

Coyle, T. R., & Pillow, D. R. (2008). SAT and ACT predict college GPA after removing "*g*." *Intelligence*, *36*, 719–729.

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: the IQ paradox resolved. *Psychological Review*, *108*, 346–369.

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

Flynn, J. R. (2012b). *Beyond patriotism: From Truman to Obama*. Exeter, UK: Imprint Academic.

Flynn, J. R. (2013). *Intelligence and human progress: The story of what was hidden in our genes*. London, UK: Elsevier.

Flynn, J. R. (2016). *Does your family make you smarter? Nature, nurture and human autonomy*. Cambridge, UK: Cambridge University Press.

Flynn, J. R., te Nijenjhuis, J., & Metzen, D. (2014). The *g* beyond Spearman's g: Flynn's paradoxes resolved using four exploratory meta-analyses. *Intelligence*, *44*, 1–10.

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

Fox, M. C., & Mitchum, A. L. (2014). Confirming the cognition of rising scores: Fox and Mitchum (2013) predicts violations of measurement invariance in series completion between age-matched cohorts. *PLoS One*, *9*(5), e95780.

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

Meisenberg, G. (2014). What are the causes of cognitive evolution? A critique and extension of psychogenetic theory. *Mankind Quarterly*, *54*, 326–328.

Nisbett, R. E. (2015). *Mindware: Tools for smart thinking*. New York, NY: Farrar, Straus and Giroux.

Ritchie, S. J., Bates, T. C., & Deary, I. J. (2015). Is education associated with improvements in general cognitive ability, or in specific skills? *Developmental Psychology*, *51*, 573–582.

Staff, R. T., Hogan, M. F., & Whalley, L. J. (2014). Ageing trajectories of fluid intelligence in late life: The influence of age, practice and childhood IQ on Raven's Progressive Matrices. *Intelligence*, *47*, 194–201.

Woodley, M. A. (2012a). A life history model of the Lynn-Flynn effect. *Personality and Individual Differences*, *53*, 152–6.

Woodley, M. A. (2012b). The social and scientific temporal correlates of genotypic intelligence and the Flynn effect. *Intelligence*, *40*, 189–204.

Woodley, M. A., Figueredo, A. J., Ross, K. C., & Brown, S. D. (2013). Four successful tests of the cognitive differentiation integration effort hypothesis. *Intelligence*, *41*, 832–842.