2011

Exploring the Flynn Effect: A Comprehensive Review of the Causal Debate

Abby J. Trimble
Claremont McKenna College

CLAREMONT McKENNA COLLEGE

EXPLORING THE FLYNN EFFECT:

A COMPREHENSIVE REVIEW OF THE CAUSAL DEBATE

SUMBITTED TO

PROFESSOR TOMOE KANAYA

AND

DEAN GREGORY HESS

BY

ABBY J. TRIMBLE

FOR

SENIOR THESIS

FALL 2011

NOVEMBER 28, 2011

Exploring the Flynn Effect:

A Comprehensive Review of the Causal Debate

Abby J. Trimble

Claremont McKenna College

TABLE OF CONTENTS

Abstract

Since its discovery in 1984, psychological investigators have continued to explore the Flynn Effect, the phenomenon of consistent and secular IQ gains within industrialized nations approximating 0.3 points per year. The most contentious debate within this field of research surrounds the purported cause of the Effect, and yet the research literature lacks a synthesis of the leading causal theories and the evidence supporting them. The principal hypothesized causal mechanisms – psychometric artifact, educational intervention, environmental changes, nutrition, genetics, gene-environment interaction model, medical improvements, and the multiplicity hypothesis – are reviewed and analyzed within the larger breadth of Flynn Effect scholarly literature. Flynn Effect causal investigation has not yielded any decisive results, and the unproductive postulation of causal theories has recently stagnated, so researchers must accept a necessary shift in the focus of their research toward a more collaborative and holistic understanding of the Effect in order to effectively determine its causes. Extensive social implications of the Effect within the scopes of special education and judicial policy necessitate the expedited revitalization of Flynn Effect research such that contemporary society may be better able to appropriately incorporate the Effect into public policy.

EXPLORING THE FLYNN EFFECT:

A COMPREHENSIVE REVIEW OF THE CAUSAL DEBATE

In 1984, political scientist James Flynn used archived American records to uncover a patterned increase in IQ scores from 1932 to 1978. The data collected from Stanford-Binet and Wechsler intelligence scales composed the results later dubbed the "Flynn Effect" by Herrnstein and Murray (1994). The compilation of Flynn's (1984) data yielded an aggregate mean IQ increase of approximately 0.3 points per year, resulting in potential gains of more than 3 IQ points per decade (Flynn, 1984). Flynn's fascination with the phenomenon continued until he determined the existence of comparable IQ gains within 14 other industrialized nations throughout the world (1987). Research has continued beyond the domain of developed countries (e.g. Bolen, Aichinger, Hall & Webster, 1995), and patterned IQ increases since have been found in 29 countries (Kanaya & Ceci, 2011).

This manifest performance improvement on common intelligence tests frequently invalidates published IQ test norms as, following the pattern of the Flynn Effect, the mean IQ test score of 100 becomes obsolete after only a decade (Kanaya, Ceci, & Scullin, 2003). In response, IQ test publishers frequently renorm their tests to better conform to well-established average score ranges. In doing so, publishers require test-takers to excel at harder questions in order to obtain the same IQ score as they would on older versions of the test. This pattern of IQ test score increase and intelligence test renorming has resulted in a rise-and-fall pattern throughout the course of an IQ test's publication tenure (Kanaya et al., 2003). The Flynn Effect, therefore, does not only detail the average increase in IQ scores within the last century, but also the decline in IQ after each renorming cycle of an IQ test (Kanaya et al., 2003). This considerable fluctuation in IQ scores suggests the Flynn Effect has the potential to be one of the

most significant psychological findings in recent history (Rodgers, 1999). Such impressive results also establish a standard of urgency within the field of Flynn Effect research: A complex psychological phenomenon with such widespread implications demands further exploration.

The search for Flynn Effect causation, to be detailed in the following chapter, has been prefaced by extensive research into the intricacies of the Effect (e.g. Teasdale & Owen, 1987; Rodgers & Wanstrom, 2007). One point of discussion surrounds the use of culture-reduced intelligence tests such as Raven's Progressive Matrices as convincing evidence for the Flynn Effect. Aggregate data collected across countries indicates the biggest IQ gains are recorded on culture-reduced tests like Raven's Progressive Matrices (Flynn, 1999). This finding is significant because some research (e.g. Jensen: see Flynn, 1999) suggests the structures of culture-reduced tests like Raven's are the most insensitive to test improvement over time.

Flynn's (1999) emphasis on the Raven's results as evidence for the Flynn Effect has become a topic of contention, as investigators (e.g. Sundet, Barlaug, & Torjussen, 2004) have used this assessment to develop theories indicating a halt or even a reversal of the Effect. Sundet, Barlaug, and Torjussen (2004) found that Flynn Effect-like IQ score gains in the latter part of the 21st century in Norway were driven almost exclusively by increases on a Raven-like test without parallel improvement on other IQ measures used in the study. The data also suggested that even the rise in the Raven-like scores stopped in the mid-1990s. These results corresponded to a recent end in Flynn Effect score increase, suggesting that the Effect, still in its academic infancy, may be empirically and sensibly obsolete (Sundet et al., 2004). This research has been extended, and further analysis has concluded that Norwegian IQ scores have even demonstrated a reversal of the Flynn Effect in the latter half of the 1990s and the beginning of the 21st century (e.g. Teasdale & Owen, 2005).

Within the greater realm of Flynn Effect research, evidence of an arrest or reversal in IQ increase trends has been treated as an anomaly limited to a specific country. For the most part, the Effect remains a pertinent topic of investigation (e.g. Rodgers, 1999), and Flynn Effect research has continued to expand its depth and reach. In favor of the existence of the Flynn Effect, more obscure, nuanced aspects of secular IQ gains have become the focus of recent research (e.g. Ang, Rodgers, & Wanstrom, 2010). One of the biggest related controversies surrounds the relationship between the Flynn Effect change in IQ scores and potential comparable gains in intelligence (Neisser, 1998). Many researchers have taken the perspective that IQ gains as measured by the Flynn Effect do not equate to real gains in intelligence (e.g. Kanaya, Ceci, & Scullin, 2003; Flynn, 1987; Flynn, 1996), and some investigators (e.g. te Nijenhuis & van der Flier, 2007) even believe the relationship between the two to be negligible. Accordingly, the current research consensus lies on the assumption that the portion of IQ gains that can be attributed to intelligence improvements is insignificant (Flynn, 1996).

Despite the questionable connection between the Flynn Effect and intelligence gains, Greenfield (1998) argues for the continued research and understanding of the phenomenon. Rather than the Flynn Effect directly reflecting upon general intelligence, Greenfield (1998) argues it is likely that the recent gains in IQ are indicative of very specific forms of intelligence increase. The Flynn Effect is representative of cognitive socialization; the rise in IQ is not a factor of increasing general intelligence, but more precisely the Flynn Effect reflects the increase in contemporary culturally-relevant intelligence. This analysis opens the door to another research theme within the broader context of the Flynn Effect: if the Effect measures culturally-relevant intelligence, investigators must determine which type of intelligence is most utilized and emphasized within modern, developed society.

The idea of differentiated forms of intelligence was proposed by Cattell (1968); in his theory, Cattell (1968) suggested two forms of intelligence: fluid and crystallized. Whereas crystallized intelligence signifies factual knowledge and comprehension, fluid intelligence is representative of problem solving ability and analytical reasoning. The study of these types of intelligence has been inherent to Flynn Effect research, and most investigators (e.g. Rodgers, 1999; Weiss, 2010; Ang et al., 2011; Zhou, Zhu, & Weiss, 2011) have determined fluid intelligence to be most sensitive to the increase in IQ scores. IQ tests reliant upon the assessment of fluid intelligence, such as the range of Wechsler intelligence scales, are more likely to reflect IQ score increases consistent with the Flynn Effect (Weiss, 2010).

Flynn Effect research literature is largely consistent in its consensus that the Effect is most closely tied to fluid intelligence, however Kanaya and Ceci (2011) wisely note that the research should not be overly specified or simplified. Echoed by Cattell's (1968) findings, the authors argue that fluid and crystallized abilities are closely tied in a reciprocal development process in which the improvement of one depends directly on the improvement of the other (Kanaya & Ceci, 2011). While Flynn Effect research more closely depends upon measurement of fluid intelligence gains, the implications of the Effect on crystallized intelligence must not be ignored.

Debates surrounding the details of the Flynn Effect continue, but the largest topic of contention centers on the search for the cause of the secular gains in IQ scores over the last century. The importance of the Effect is no clearer than in the causal debate surrounding it, for some researchers (e.g. Flynn, 1999) believe the true meaning and implications of the Effect remain unknown until its cause is determined.

Flynn Effect Causal Theories

It is true that investigators have endeavored to further the field's understanding of the Flynn Effect, yet the literature about the Effect still lacks a comprehensive synthesis of causal explanations. The following section will detail and critique several of the proposed causal mechanisms within the context of the larger breadth of literature.

Psychometric Artifact

Recent Flynn Effect research has shown wide-spread cumulative support for the existence of enduring, secular gains in IQ scores within developed countries (e.g. Daley et al., 2003; Nettelbeck & Wilson, 2004), and yet some investigators (e.g. Brand, Freshwater, & Dockrell, 1989; Wicherts, Dolan, Hessen, Oosterveld, van Baal, Boomsma, et al., 2004) do not believe these increases are indicative of anything other than psychometric or testing artifacts. One common critique of Flynn Effect findings is that the tests used to assess IQ, especially those from the Wechsler Intelligence Scales, lack measurement invariance; consistent gains in IQ cannot be attributed to an increase in IQ-related variables that the IQ tests purport to measure (e.g. Wicherts et al., 2004). Still other researchers (e.g. Brand et al., 1989) question the validity of contemporary IQ measures because of their overreliance on evidence for intelligent guessing as an indicator of higher IQ. Critical researchers like Brand et al. (1989) assert that IQ tests reward flexible, creative test takers over conscientious ones, and this discrepancy creates an artificial and imperfect understanding of IQ.

Beaujean and Osterlind (2008) designed a study to test explicitly the artifact hypothesis. They analyzed scores obtained from the Children of the National Longitudinal Survey of Youth (NLSYC) to determine if the data did indeed reflect real patterns of the Flynn Effect as opposed to testing artifacts. The authors assert that previous analysis of the NLSYC data to determine the

existence of the Flynn Effect relied upon Classical Test Theory (CTT) to structure the investigation; this methodology, though, is inherently incapable of analytically separating any psychometric artifact from a real effect (Beaujean & Osterlind, 2008). In order to differentiate between genuine increases in IQ and the existence of a psychometric artifact, the authors utilized analysis based on Item Response Theory (IRT) rather than CTT. The use of IRT over CTT is advantageous in the case of Flynn Effect research because it allows the investigators to separate the constructs IQ tests are designed to measure and the scores that suggest to measure them (Beaujean & Osterlind, 2008).

Upon the interpretation of the IRT-based analysis of the results of the study, Beaujean and Osterlind (2008) found a much smaller increase in IQ test scores over time than that reported by the Flynn Effect; in the case of the Peabody Picture Vocabulary Test-Revised, the increase in IQ scores was insignificant. The authors of the study argue that these results cast doubt upon previous literature implementing CTT-derived scores to verify and confirm the existence and pattern of the Flynn Effect. While careful consideration of the analytical methodology used to assess the Flynn Effect is necessary and should be respected, the results of this study are exploratory rather than definitive. The conclusions should be used to promote further study of the subject of alternative methods of data analysis in Flynn Effect research.

Several investigations have provided disconfirming evidence for the Flynn Effect in the vein of the psychometric artifact hypothesis, the purpose of which is to ultimately question the existence of the Effect. While these research reviews are important to consider, they also must not be taken as fact but rather contextualized within the broad research literature of the Flynn Effect. The search for disconfirming evidence, as echoed by Rodgers (1999), is an important part of the scientific process, and theories cannot be developed without proper questioning.

The arguments surrounding the psychometric artifact hypothesis are especially compelling in that data collection and analysis form the foundation for all Flynn Effect theories. Without IQ test taking and score analysis, Flynn and subsequent researchers (e.g. Daley et al., 2003; Rodgers & Wanstrom, 2007) could not have proposed and supported the patterned, secular increase that is the Flynn Effect. Accordingly, testing and data analysis processes critiqued by the psychometric artifact hypothesis are intimately tied with this IQ phenomenon.

The most fundamental evidence contrary to the artifact hypothesis is the widespread research supportive of Flynn's original findings. The Flynn Effect has been identified in IQ scores representative of as many as 29 countries (Kanaya & Ceci, 2011), with the supportive research now extending to rural and underdeveloped countries as well (Daley et al., 2003). It may be the case that Flynn Effect researchers are too far engaged in Rogers' (1999) cycle of the mistaken search for corroborating evidence, however it seems unlikely that so many separate cases of confirmatory support of Flynn Effect would be found around the world if that were the case. Evidence authenticating the psychometric artifact as a Flynn Effect causal theory provides, if nothing else, justification for Flynn Effect to be observed and scrutinized within a larger analytical context. For example, many researchers (e.g. Teasdale & Owen, 2005; Kaufman, 2010) critically focus on the use of scores from Raven's Cultural Matrices as evidence for the Flynn Effect because Flynn (1984) determined it to be the IQ test most sensitive to the Effect. Instead of perpetuating the cyclical pattern of confirming evidence, advocates of the artifact hypothesis should take steps to understand why Raven's provides the most reliable supportive evidence for the Flynn Effect.

The field of Flynn Effect research is stagnating because investigators prefer critiquing existing theories over proposing new evidence; advocates of the psychometric artifact hypothesis

must propose constructive ways in which the Flynn Effect is wrongly represented within the current research. By pinpointing specific flawed aspects of IQ tests utilized for Flynn Effect support, the psychometric artifact hypothesis will gain more credibility within the field.

Educational Intervention

Broadly stated, the topic of education is obviously related to the understanding of the Flynn Effect in that the goal of education is to enable students to analyze and criticize, both of which are cognitive processes undoubtedly related to IQ (Flynn, 1998). Accordingly, in the search for causal explanations of the Flynn Effect, many investigators have directed their focus upon the realm of education (e.g. Williams, 1998; Blair, Gamson, Thorne, & Baker, 2005). Instead of a concerted focus on one aspect of the educational system, though, education hypotheses encompass a wide-array of educational improvements and interventions.

One such explanation suggests increased access to education as a fundamental factor in the IQ increases contributing to the Flynn Effect. During the 20th century, developed countries became more urbanized, allowing the expansion of semiformal and formal educational institutions (Williams, 1998). As education became more socially relevant to the social and intellectual development of children and teens, individual commitment to education increased. This phenomenon is reflected in each of the countries from which Flynn collected data in that each country demonstrated records of its citizens spending larger fractions of their lives in school (Flynn, 1998). Williams (1998) echoes the increasing prevalence of education within the lives of American students, noting that the mean number of years of educational attainment in the 1990s was four to five years longer than that of the 1930s. This data implies that the people taking IQ tests in the early 20th century had been exposed to many fewer years of schooling than

contemporary test-takers, thus likely producing patterns of Flynn Effect-like IQ score augmentation.

A link that strengthens the access to education hypothesis is the research exhibiting education-based increases in crystallized and fluid intelligence (Kanaya & Ceci, 2011). Fluid intelligence, as mentioned above, has been identified as the type of intelligence most sensitive to the Flynn Effect (Zhou, Zhu, & Weiss, 2010; Ang et al., 2010). Accordingly, the integration of these to factual findings indicates that increased involvement in school will result in higher IQ scores for students.

A similar, yet differentiated explanation for the Flynn Effect lies in the overall increase in the availability of and access to preschool education. In the case of Teasdale and Berlinder (1991), their research review focused primarily on kindergarten as a form of preschool education. The study concluded that access to kindergarten education will likely result in the improvement of several intelligence constructs, purportedly causing enhanced performance on IQ tests (Teasdale & Berlinder, 1991). Essentially, the results indicate that adults with access to kindergarten at the beginning of their academic careers are likely to have higher educational levels and test scores, and these two factors are typically tied with higher levels of IQ. This research is important because of the recent push to encourage children's participation in preschool and kindergarten classes within developed countries like the United States (Kirp, 2007). In addition to many other short-term benefits of preschool education, such as early literacy and school readiness (Barnett, Brown, & Shore, 2004), access to preschool and kindergarten education may implicitly cause an increase of student scores on IQ tests. Thus the Flynn Effect may be partially tied to the increasing prevalence of kindergarten education within industrialized countries.

Concurrent with the trend toward increasing educational access, recent improvements made to the contemporary academic curriculum contextualize another branch of educationally-based Flynn Effect causal theories. The development of new curricula, with a specific focus on math instruction, is believed by many researchers (e.g. Blair et al., 2005) to constitute at least a portion of secular increase in IQ (Beaujean & Osterlind, 2008). One study focused on the rising American population mean IQ concludes that, in addition to the population's increasing access to schooling, the increasing cognitive demand of the mathematical curriculum is a likely cause of environmentally-driven gains in intelligence between generations (Blair et al., 2005). According to the researchers, because the increased difficulty of the math curriculum affects students at an increasingly early age, there continue to be substantial intergenerational differences in IQ as reflected by the Flynn Effect.

Differences in the contemporary math curriculum have also been attributed to the recent and mounting focus on activity-based learning rather than traditional memorization or individual calculation. Williams (1998) asserts that the modern school shift toward activity-based representations of math concepts constitutes a more global change in instructional focus from fact memorization toward problem-solving skills training. Problem solving practice is intimately tied to the sharpening of fluid intelligence, so the theory states that change of math curricula towards a more fluid approach underscores contemporary students' higher IQ scores (Williams, 1998). Moreover, it is unlikely that the teaching style dedicated to the instruction of math has so singularly and uniquely diverged from the instructional process as a whole. As asserted by Flynn (1998), there exists evidence that schools in general are teaching better overall problem-solving skills, training that can be directly applied to the context of IQ tests. The proliferation of interactive learning tools is likely found within the entire academic curriculum, though perhaps

the shift to this type of instruction has been more nuanced in other subjects than in the case of mathematics and for this reason it has been noted within the context of Flynn Effect causal hypotheses. While math instruction may be one cause of Flynn Effect IQ gains, it is more likely that an entire curriculum shift toward interactive instruction and problem-solving practice plays a part in the patterned increase of IQ scores across industrialized countries.

Another prevalent phenomenon often associated with the development and alteration of school curricula, especially within the United States, is the tendency to teach to the test (Popham, 2001). In modern society, standardized tests are relied upon to gauge the progress of individual students, teachers, schools, districts, and educational systems on the whole (Urdan & Paris, 1994). Keeping aside the important relationship between standardized test performance and federal funding, much of a student's personal potential is dependent upon his ability to excel on standardized tests (Pritchard & Wilson, 2003). For these reasons, contemporary public education caters to students' eventual ability to excel on these tests.

Williams (1998) argues that teachers' matching of instruction to the demands of different standardized test has significant effects on the overall tendencies of IQ test scores of students in developed countries. Methods used to prepare students for these tests include both fact drilling and explicit, lengthy training of test-taking skills, activities that require both crystallized and fluid intelligence, respectively (Williams, 1998). Increases in fluid intelligence resulting from this type of test-taking preparation would likely be reflected on fluid intelligence-sensitive IQ test scores, ultimately resulting in patterns of the Flynn Effect. This theory is notable simply because many advocates of educational reform argue that the strategy of teaching to the test negatively affects a student's ability to learn within the classroom (Firestone, Schorr, & Monfils, 2004). If Williams' (1998) assertion is valid, however, it may be noted that this instructional

method not only increases students' performance on all-important standardized tests but also reflects positively in their IQ.

The link between rigorous preparation targeting improved standardized test-performance and increasing IQ test scores is neither complex nor unbelievable. The point where the teaching to the test hypothesis wavers, however, is in the IQ test improvement as evidence for overall intelligence improvement. While it may reflect positively upon college readiness or IQ test scores, preparation for standardized tests within the classroom is not necessarily pertinent to real-world applications of intelligence until later studies can validate the strength of this relationship.

Intimately tied to the hypothesis of IQ increases as caused by standardized test preparation is the causal hypothesis postulating that increased educational funding has contributed to the development of better schools, which produce smarter, more capable students who are more likely to excel on IQ tests. Specifically, funding theoretically enables more quality educational institutions to incorporate teacher-training programs that improve the educational experience of the student. Williams (1998) provides data-based evidence for the increase in educational funding during the past half-century: the total real increase in educational spending per pupil was 61 percent between the years of 1967 and 1991 (p. 135). As much of this money has been dedicated to the augmentation of teachers' salaries (Ingersoll, 2002), this theory cannot decisively conclude that the increase in funding has directly contributed to the improved quality of education for the average child within a developed country. However, it seems plausible that the increase in funding has contributed to the Flynn Effect in some way, because, ideally, the augmented funding has improved the state of education for all children, a progression which theoretically yields an increase of average student performance intelligence testing. At the very least, funding is likely to result in higher teacher salaries which are likely to have a positive

effect on the learning within the classroom; higher salaries are likely to yield higher teacher dedication to the profession (Billingsley & Cross, 1992), which may result in more creative or successful instructional methods.

Empirical research conducted by Ang et al. (2010) further supports the funding-based improvement of the educational system as a causal means for the increase in IQ scores. The study was conducted through the use of math scores obtained from the NLSYC data as a means to test the pattern of the Flynn Effect within demographic subgroups. Through the analysis of these subjects' longitudinal PIAT-Math scores, Ang et al. (2010) concluded no significant subgroups difference; instead all subgroups confirmed to the proposed aggregated norm of an approximate increase of 0.3 IQ points per year (Flynn, 1984). Accordingly, the research conclusions reflected that any potential cause of the Flynn Effect must be independent of gender, race/ethnicity, maternal education, household income, and urbanization. The educational improvement hypothesis, therefore, is supported by the results of the study. According to the authors, the educational system is characterized by a manifest effort to improve all instructional aspects in order to better support student learning (Ang et al., 2010). The researchers clarify, though, that the likelihood that this type of improvement is as consistent as the Flynn Effect is unlikely; however their results are evidence for funding-based educational improvements as one causal explanation for the Flynn Effect phenomenon.

While the education funding hypothesis is cogent on its face, the United States No Child Left Behind Act (NCLB) of 2001 is the source of a potential critique. NCLB is structured as a federal funding program contingent upon student attainment of standardized test score standards (NCLB, 2002). Accordingly, the Act introduces higher levels of both federal funding and standardized test taking into American public schools, both of which encompass causal

hypotheses for the Flynn Effect. Following the logic of the federal funding and standardized testing hypotheses, then, the introduction of NCLB she be followed by a sharp increase in IQ score gains, positively deviating from the aggregate Flynn Effect mean of 0.3 points per year. Whereas the educational environment prior to NCLB included both federal funding and standardized testing, the implementation of the Act guaranteed that the two phenomena were profligate within the academic context. Any effect that funding and testing have on IQ, then, should be multiplied by the augmentation of their involvement in public education. The federal funding and standardized testing hypotheses fail, however, because the Flynn Effect has remained somewhat constant since NCLB's introduction in 2001; some researchers (e.g. Sundet, Barlaug, & Torjussen, 2004; Teasdale & Owen, 2005) even argue that the Flynn Effect is no longer operative within this contemporary timeframe.

This legislative measure, thus, provides disconfirming evidence for the educational funding and standardized testing Flynn Effect hypotheses because NCLB has not yielded tremendous gains in IQ scores. Instead of the proposed causal relationship between federal funding and IQ score improvement, it is better asserted that federal funding and related academic interventions, like standardized testing, provide students with an opportunity to improve academically. This improvement may be reflected on their enhanced ability to perform on IQ tests, but educational funding does not necessitate this improvement.

Other researchers (e.g. Schooler, 1998) have championed even more specified education-related Flynn Effect causal explanations, citing particular aspects of the educational experience implicit in the trend toward increased IQ scores. Despite the proliferation of educational factors within the context of Flynn Effect explanations though, the divided focus of education hypotheses on distinct, small-scale educational aspects is ill-advised and unnecessary. Education

intervention should be considered holistically when used to explain the Flynn Effect because the likelihood that one limited educational factor constitutes the source of the Effect is small at best. Accordingly, the complete experience of contemporary education in developed countries likely constitutes one component in the social equation that determines the pattern of the Flynn Effect.

Environmental Change

The structure of the contemporary environment is the source of many Flynn Effect causal theories. As compared with social settings in the past, the environmental changes hypothesis argues that more advanced cognitive skills are required to navigate the complex environment, and these cognitive skills are indicative of the fluid intelligence gains represented in the Flynn Effect. Essentially, these theories are consistent in their assertion that the conditions of modern environment are distinctive because they require the mastery of fluid intelligence in order to successfully navigate the modern world.

Greenfield's (1998) understanding of the environmental condition is entirely relevant to Flynn Effect research. Consistent with the synthesis provided by Ang et al. (2010), Greenfield has ascribed the Flynn Effect to massive environmental changes including the proliferation of technological entertainment and communication media such as movies, video games, and computers. In Greenfield's (1998) view, the patterned increase in IQ is representative of an evolution of cultural intelligence through which people in developed countries are gradually prioritizing iconic images over written ones. The abundance of visual mediums such as television underscores Greenfield's proposed cultural shift, and it is possible that the modern transition to iconographic representation has facilitated the consistent rise in nonverbal IQ reflected in the Flynn Effect. Many IQ tests (e.g. Wechsler intelligence scales) use measures of nonverbal, iconographic representation to assess IQ, so it is logical that people with increased

accessibility to and practice with iconographic images will achieve higher IQ scores. The theory extends the logic of many subsets of the educational hypothesis by asserting that, in addition to the confines of the classroom, the environment at large requires the frequent use and practice of fluid intelligence. Schooler (1998) argues that this cultural transformation began during the industrial revolution, and since that time there has been a steady introduction of more mediums demanding fluid intelligence, which has essentially created a fundamentally more complex environment. Schooler (1998), in concordance with Flynn (1998), asserts that gradually increasing environmental complexity is conducive to a parallel increase in intellectual functioning due to amplified cognitive demands of daily life.

The process of industrialization as manifested within technological development is largely restricted to developed countries because the source of environmental complexity is coupled with economic complexity (Schooler, 1998). Accordingly, populations exposed to overall social complexity should exhibit related intellectual gains. The structure of Schooler's (1998) hypothesis expands upon individual interactions with the environment to include the relevance of broader societal structures as a potential cause of the Flynn Effect. Recent research conducted by Wei and Putallaz (in press) confirms the proliferation of environmental complexity, or environmental stimulation, as a potential source of the Flynn Effect. The study endorses video games and computers as sources of such stimulation that contribute to rising IQ within the whole population, including people in the upper stratums of intelligence.

The complete explanatory environmental hypothesis is perhaps a more compelling individual causal explanation than any other. Proliferation of requisite problem-solving skills for the comprehension of the progressively more complex modern environment is a convincingly thorough and practical explanation. Whereas the other theories unnecessarily focus on a portion

of modern culture as the cause of the Flynn Effect, the environmental explanation encapsulates an entire phenomenon of developed society. That phenomenon, the progression of contemporary industrialized culture to promote the acquisition of fluid intelligence, is verified in numerous contemporary contexts, including children's activities on the sides of McDonald's Happy Meals (Williams, 1998). Accordingly, denial of some environmental aspect in the causal equation of the Flynn Effect is futile; the holistic environmental conception of the Flynn Effect has marked strength in its overt and commonsensical applications to modern society.

Nutrition

Nutritional explanations have provided the basis for a class of popular theories, and it has been argued that contemporary dramatic changes in available nutrition have been responsible for some part of the IQ increase (Ang et al., 2010; Beaujean & Osterlind, 2008). Lynn (1987; 1989) is a devote advocate of the nutritional hypothesis. His research concludes that a substantial portion of the global IQ increase can be explained through the nutrition hypothesis which is corroborated through the analysis of parallel gains in height and head size of similar magnitude. Lynn (1989) finds positive correlations between the three variables – IQ, head size, and height – and asserts that increase in height and head size are evidence for the existence of parallel gains in intelligence.

Lynn (1990) is of the complete conviction that IQ gains were representative of real intelligence gains, and his evidence of secular increases in height, head circumference, and brain size served to argue that neurological development and the functioning of the brain were improving with increased access to nutrition. He argues that the link between height and access to nutrition is irrefutable and evidenced through the parallel increases in height and improvements in the standard of living in developed countries. Lynn (1990) links neurological

development to height, arguing that the same nutritional improvements have led to increased brain size and evidence of improved neurological functioning. The theory follows the logic that larger heads, holding larger brains, are necessarily more intelligent than smaller ones with smaller brains; this logic has been corroborated by more than ten studies verifying a corresponding positive association (Lynn, 1990). Lynn's generous, and perhaps erroneous, prediction of the true correlation between head size and intelligence serves as his primary evidence for the nutrition hypothesis.

Lynn's theoretical contributions have been supported many times through the work of other Flynn researchers (e.g. Eysenck & Schoenthaler, 1997). Martorell (1998) argues for height as an indicator of nutrition through an analysis of height trends in developed countries. He concludes that the evidence linking poor nutrition to impaired cognitive growth in developing countries suggests an irrefutable link between the remarkable nutritional improvements of the 19th and 20th centuries and the persistent improvement of IQ scores in those countries. Eysenck and Schoenthaler (1997) extend these results by arguing that vitamin and mineral supplementation for all children would result in enhanced social benefits and personal intellectual improvement. Consequently, the authors believe the possibility that IQ can be raised through vitamin and mineral supplementation.

Flynn (1999) requested that causal investigators address the complex details of the Effect in the formation of hypothetical causal mechanisms. The nutrition hypothesis moderately addresses two of the three most perplexing Flynn Effect phenomena: the consistent, stable rate of IQ gain and the lack of intragenerational IQ differences throughout the period of secular IQ increase. The increasing availability of complete nutritional supplementation in developed countries serves to provide at least some explanation for the generational phenomenon of the

Flynn Effect. The availability of nutritional food has changed substantially in industrialized countries, and so the standards of consumption have also changed (Delgado, 2006). Whereas older generations had access to perhaps more wholesome food options, recent generations are exposed to copious options, many of them processed or preserved (Striffler, 2005). As one develops, food consumption standards form and solidify, creating intragenerational food customs. This basic explanation serves the problem of generational differences within the supporting Flynn Effect research: speaking generally, average people within a generation have relatively similar eating habits, while those standards can vary widely from the habits of previous or subsequent generations.

Additionally, the nutrition hypothesis provides some insight into the explanation of the Effect's patterned consistency across countries and socioeconomic statuses. While food customs and norms are notably different across country lines, it is without question that developed countries have increased access to more food and, in many cases, more nutritional food (Delgado, 2006). The recent phenomenological standard of processing food has reached widespread levels, certainly altering multiple countries' access to nutrition (Popkin, 2006). This process is one of a few notable and measurable world-wide developments that are potentially related to the Flynn Effect, and for that reason much postulation and hypothesizing has been done surrounding this explanation.

The validity of nutritional hypotheses could be easily tested through the introduction of nutritional supplementation in underdeveloped countries, and yet some researchers are unconvinced of causal legitimacy. Some (e.g. Ang et al., 2010) suggest that the nutritional hypothesis only serves within the larger context of a multiply-determined Flynn Effect. Ang et al. (2010) found results consistent with the nutrition hypothesis, and yet the authors remained

unconvinced of the nutrition hypothesis as a compelling single explanation. Some investigators (e.g. Flynn, 1999) provide even more comprehensive critiques of the nutrition hypothesis.

Flynn's (1999) basic criticism of the nutrition hypothesis concludes that nutritional improvement cannot be as prolonged and consistent to reflect in the patterned Flynn Effect. Flynn's critique is supported by a study indicating that the effects of vitamin and mineral supplementation in California on IQ were only significant with the introduction of a moderate supplement, whereas small and modest supplements have insignificant effects (Schoenthaler, Amos, Eysenck, Pertiz, & Yudkin, 1991 as cited in Flynn, 1999). How is it possible, Flynn (1999) argues, that so many nations would maintain perfected and consistent nutritional improvements throughout recently history to result in the most significant IQ gains? By all accounts, this type of patterned nutritional improvement seems unlikely.

While Flynn's (1999) critiques of the nutrition hypothesis are sound, the relevance of nutrition within the broad scope of the Flynn Effect is irrefutable. The international phenomenon of intergenerational differences of the Flynn Effect is hard to explain without some tie to global improvements in nutrition and access to food within industrialized countries because few other societal changes have been as widespread and prevalent across the globe. Accordingly, Lynn's (1990) theoretical basis for the nutrition hypothesis is convincing because of the anecdotal and fundamental ties between cognitive functioning and nutritional intake. It is almost obvious that improved nutrition would result in mirrored improved cognition as Lynn (1990) suggests, so his assertion of nutrition as a more powerful determinant of intelligence appears sound.

Evidence for the nutrition hypothesis, however, is not strong enough to support Lynn's (1990) conviction that improvements in nutrition constitute the fundamental cause of the Flynn Effect. While nutrition may be a stronger determinant of improved IQ than previously

suggested, such complete and conclusive data does not exist to eliminate all other explanations of the Effect. In other words, wide-reaching evidence for the Flynn Effect across gender, race, socioeconomic status, and urbanization suggest that the nutrition hypothesis plays at least some role in the Effect, but improved nutrition does not yet provide a complete explanation of the patterned increase in IQ.

Genetics

The genetic explanation of the Flynn Effect is most often framed in the context of heterosis, or biological enhancement through outbreeding. Under the basic assumption of heterosis, animals will tend to develop improved or increased functioning as mating pools expand because natural selection favors more advantageous biological qualities, including intelligence. Mingroni (2007) is considered the pioneer of the heterosis hypothesis within the context of the Flynn Effect, and he proposes that the Flynn Effect is caused by a broad increase in geographic breadth of the mating pool due to the availability of travel within developed countries; this greater mating pool should lead to a positive selection pressure on intelligence and other biological characteristics consistent with IQ increase (Mingroni, 2007).

Whereas most causal explanations of the Flynn Effect are purely environmental in origin, Mingroni's (2007) heterosis hypothesis is unique and thematically relevant. Widely accepted estimates of the heritability of IQ significantly eliminate the existence of substantial environmental effects on IQ variability (Neisser et al., 1996; Jensen, 1998; Herrnstein & Murray, 1994). Accordingly, Mingroni (2007) asserts that the heterosis explanation provides the only meaningful causal justification of the Flynn Effect. Environmental factors can not be the accurate Flynn Effect causal predictors due to high heritability and low environmentally-based IQ flexibility; such high heritability estimates necessitate a genetic theory (Mingroni, 2007). The

heterosis hypothesis also asserts that many other heritable traits have increased secularly along with the patterns of the Flynn Effect (Mingroni, 2004). The only plausible explanation of these parallel phenomena is the existence of a broad-based genetic change like heterosis (Mingroni, 2004).

While the theoretical reasoning behind the heterosis hypothesis is strong, there is insufficient evidence to suggest an expansion of the genetic pool dramatic enough to create real hereditary change. Additionally, based on widely established social psychological patterns of ingroup preference, the odds are unfavorable that enough people would choose to mate outside their familiar genetic pool such that a global phenomenon would occur. Finally, one of Mingroni's (2007) fundamental supporting arguments is that environmental effects cannot be large enough to affect IQ given high IQ heritability, however, following the Dickens and Flynn (2001) genetic-environmental interaction model, the existence of high IQ heritability does not eliminate environmental intervention as a cause of IQ change.

Gene-Environment Interaction Model

Authors Dickens and Flynn (2001) formulated the causal gene-environment interaction model to address the paradox between reported significant genetic heritability estimates and large environmentally-caused differences in IQ. As a consequence of the approximated heritability of IQ (Neisser et al., 1996), environmental variance theoretically accounts for little variance in adult IQ, and yet the Flynn Effect is one documented phenomenon that insinuates large-scale environmental effects on individual IQ.

In order to explain this paradox, the authors asserted a reciprocal link between IQ and environment (Dickens & Flynn, 2001). Through the analysis of their theory, Dickens and Flynn (2001) conclude that the reciprocal causation model of genes and environment could potentially

multiply environmental effects such that small changes in the environment can produce substantial changes in IQ. The gene-interaction model, therefore, casts doubt on Mingroni's (2004; 2007) heterosis hypothesis as genetics are not the only feasible predictor or regulator of IQ. Instead, environmental factors can rival genes in their importance, especially when the effect of environmental change is permanent or widespread.

Medical Improvements

Advocates of the medical improvements causal explanation suggest that increased access to better medical care, a conditional characteristic of developed countries (Mosley & Chen, 1984), contributes to the overall increase of IQ. Steen's (2009) rising tide hypothesis asserts that the cause of the Flynn Effect lies in patterned medical improvements as evidenced by the coupling of parallel IQ and medical procedural improvements. Through the process of medical advancement, physicians were able to eradicate illnesses and other medical conditions that dampened intelligence in the past (Steen, 2009). Further evidence for this theory indicates that declines in serious disease and illness-causing conditions correlate strongly with continued cognitive gains, indicating some link between public health and aggregated IQ (Steen, 2009).

The research conducted by Ang et al. (2010) addressed the relevance of the medical improvements hypothesis within the context of the results of the study. As with the other supported causal explanations, the existence of medical improvements must occur across all racial classes and urban categories in order to be consistent with the lack of Flynn Effect subgroup differences. Certainly, examples of such medical intervention exist within the modern and developed world; two such examples are the treatment of tap water with fluoride to prevent tooth decay and the widespread elimination of lead-based paint (Ang et al., 2010). The comprehensiveness of this theory, however, to explain the persistent increase of IQ is

questionable at best as it is unclear that far-reaching public health interventions with a restricted medical scope are capable of such great increases of IQ.

Instead of medical improvements constituting a causal theory for the pattern of the Flynn Effect, these medical improvements are likely indicators of improved standards of living characteristic of developed countries. In addition to improved access to high-quality medical care, citizens of developed countries also experience greater access to educational and nutritional resources, both of which are factors also implicated in the causal debate of the Flynn Effect. Accordingly, rather than a singular explanatory tool, the medical improvements hypothesis is better served as a piece within the larger context of environmental change resulting in the Flynn Effect.

The Multiplicity Hypothesis

The basic assertion of the multiplicity hypothesis, pioneered by Jensen (1998), is that many small factors, some likely unspecified, have combined to create the Flynn Effect. Wei & Putallaz (in press) characterize the multiplicity hypothesis by four main components, each of which is largely representative of another theoretical explanation mentioned above. First, the hypothesis argues some effect of the broad use of standardized tests, resulting in widespread familiarity with test taking like that of the IQ test. The relationship between standardized testing and the Flynn Effect has already been discussed at length, however the multiplicity hypothesis offers a different perspective. The practice effect achieved by the systematic repetition of standardized tests throughout the academic career primes contemporary students for improved IQ test scores simply because they are better prepared to take the test.

Jensen's (1998) second factor that operates to create the Flynn Effect is the broad trend of educational improvements, including increased access and longevity of academic dedication.

Simply put, more people are going to school for more years within industrialized countries. The discussion about the educational hypothesis, also above, provides substantial insight into the logic of this portion of the multiplicity effect. Education is intimately tied to IQ, so it logically follows that increased access to more years of school would result in an aggregated mean increase in IQ scores (Williams, 1998). Additionally, the contemporary academic requirement of complex cognitions enables students to better perform on IQ tests (Blair et al., 2005).

The third factor of the multiplicity hypothesis is the trend toward improved nutrition and healthcare in developed countries. Without ample nutrition and proper healthcare, people will not have the time or the resources to dedicate their energy to improving their IQ; additionally, these improved standards of living are indicative of a more complete environmental complexity in which the acquisition and grooming of fluid intelligence skills are encouraged.

Finally, the multiplicity hypothesis suggests that advances in obstetrical practices and preventative inoculations protect most of the developed world from formerly common childhood diseases. Tied to the improvement of nutrition and healthcare, this factor suggests that developing countries are now able to eliminate most, if not all, childhood obstacles that previously impeded access to education and development of intellect.

As is clear from the description of the hypothesis, the multiplicity causal explanation of the Flynn Effect is attractive in that it provides an explanation for the consistency of the Flynn Effect across time and across culture (Ang et al., 2010). If only one factor were to account for the Flynn Effect, its direction and magnitude could seemingly fluctuate between different countries, but the consistency of the Effect suggests a balance of multiple factors, yielding a constant pattern of IQ score increase.

The reliability of the multiplicity hypothesis to explain the regular pattern of IQ score increases is compelling, yet not entirely convincing. While educational effects and standardized test practice constitute a strong causal component, the literature and research are not conclusive enough to argue that improved nutrition, healthcare, and disease prevention make up a large portion of IQ increase. Additionally, in most Western, developed countries, early childhood illness and fatality have been eliminated for some time (Rosano, Botto, Botting, & Mastroiacovo, 2000), so it is unlikely that the intervention of modern inoculation and obstetrical practices would still contribute so substantially to the patterned IQ increases.

Formulation of Flynn Effect Causal Theories

Despite its relatively young empirical existence, the Flynn Effect has already been widely discussed within the discipline of psychology (Rodgers, 1999). While a "massive" change in IQ poses significant consequences within the field, the realm of IQ extends far beyond psychological exploration (Flynn, 1984). For that reason, numerous researchers have focused on the causes of the Effect, with the hope of better understanding its theoretical and practical import. As with most aspects of the Flynn Effect, though, controversy surrounds the postulation of theoretical causes of the Effect. More explicitly, there exists discontent in the field regarding the proposition of multiple theories without a proper broad synthesis of the completed research. In other words, it seems right now as if too many people are concerned with simply making their ideas known rather than presenting meaningful answers as insight into potential cause of the Effect (Weiss, 2010).

Some psychologists (e.g. Weiss, 2010; Rodgers, 1999) assert that many Flynn Effect investigations have been structured in order to support or confirm prefabricated hypotheses. Weiss (2010) claims that the field of Flynn Effect research is crowded with investigators who

mine for idiosyncratic, confirming data points around which popular theories are formed. Researchers should, however, be much more concerned with identifying empirical evidence that should not be present if the hypothesis is true (Weiss, 2010). Only with sound methodology and objective analysis can the hidden cause and ultimate significance of the Flynn Effect come to light.

Weiss' (2010) argument is extraordinarily compelling, especially as the broad implications of the Flynn Effect are highlighted frequently in contemporary psychological journals: Instead of a systematically sustained focus on the assertion of new causal theories, researchers should redirect their concentration to finding disconfirming evidence for current theories. Rodgers (1999) argues that the field is best advanced when contrasting causal mechanisms are compared and systematically tested. Current researchers, though, seem more preoccupied with their own isolated conjecture rather than placing empirical results within the complete context of the Flynn Effect. To the detriment of their research, Flynn Effect investigators have been caught up in the hype surrounding the Effect, leading them to forsake their empirical training so that they may contribute to the armchair causal debate.

Not only have the methodology and rationale surrounding current causal theories of the Flynn Effect been critiqued, but some researchers go so far as to evaluate even the practice of causal speculation. Rodgers (1999) suggests that Flynn Effect researchers have not sufficiently questioned Flynn's (1984; 1987) methodology that led to the formation and verification of the theoretical Effect. Without this essential methodological step, the theory is not sound enough to warrant such concern with potential causes.

Consistent with Weiss (2010), Rodgers' (1999) argument suggests that Flynn Effect researchers have become too supportive and accommodating in their empirical approach. While

the Flynn Effect could be one of the most influential findings in contemporary psychological research (Rodgers, 1999), its purported significance necessitates that the theory's methodology should be verified before its widespread and complete acceptance. The Flynn Effect's potential applied significance in the field should, in fact, encourage researchers to question the methodology and research behind it. Without proper critique, the real hidden meaning and causes of the Effect could be lost.

The field still lacks a complete understanding of the nature of the Effect, and for that reason causal postulation has yet no meaningful place. Instead of such a global focus on the causes of the Effect, researchers must focus their efforts to better understand its meaning. Rodgers' (1999) argument is fundamentally persuasive and transparently basic. Ideally, the field would not require an investigator to suggest that researchers explore the fundamental logic behind a theory before rushing to explain its causes and implications. In this case, however, the backwards research of the Flynn Effect seems to be the norm.

Flynn (1999) has a different approach to the formation and promotion of numerous causal explanations of his Effect. In contrast to Rodgers, Flynn (1999) argues that the search for causal answers is important and relevant; this divergence in conviction is not surprising as it goes without saying that Flynn believes his Effect to be valid and factual. His critique of the hypothesizing of causal mechanisms lies in his understanding of what these postulations achieve. Flynn (1999) insists that the hypotheses offered up until this point, such as the nutrition and education hypotheses, are not appropriate in that they do not address the full complexity of the Effect, so they are not apt causal theories.

Flynn's advocacy for the advancement of hypothetical causal mechanisms has taken hold of the current research imperative, as recent investigators have once again embraced their role as

causal postulators (e.g. Mingroni, 2007). With respect to Rodgers (1999) and Weiss (2010), though, most contemporary publications attempt to consider the relevance of suggested causes within the context of the "black box"[1] of the Flynn Effect, rather than simply asserting new ideas.

Although his article was published more than a decade ago, it seems as though Rodgers still has it right. A wide range of causal theories was proposed in 1998 with the publication of *The Rising Curve* (Neisser), and yet it does not seem that any of those theories has gained more credibility than any other. Flynn Effect researchers have performed studies and published commentaries advocating for one causal theory over another, however, no conclusive evidence has been found to shift the focus of the field in one causal direction.

After so many years of research, it is clear that the understanding of the Effect is limited. In order to truly understand the implications of the Flynn Effect, the field must continue to explore the cause in the face of this intellectual stagnation. Perhaps this lack of progress, instead of being frustrating, is indicative of a necessary shift in the approach to the Flynn Effect causal debate. This progressive cycle of unknowing will undermine the eventual understanding of the Flynn Effect unless researchers can rebuild their understanding of progress as it applies to the cause of the Effect. Perhaps contemporary investigation is missing the point: the lack of progress in the process of causal discovery is indicative of the real Flynn Effect cause: multiplicity. Following this logic, the multiplicity hypothesis has been unintentionally supported through the relative stagnation of specified causal theory-building. It is nearly impossible to encapsulate the cause of the Flynn Effect within one theory if indeed that effect is multiply

[1] The "black box" behind the Flynn Effect relates to the debate surrounding the causal explanation of the Flynn Effect. As previously detailed, investigators have not yet been able to identify a concrete environmental, social, nutritional, or genetic causal mechanism to completely explain the Flynn Effect. Therefore, researchers cannot yet determine whether the Flynn Effect represents real IQ and intelligence gains or whether the Effect indicates a simple psychometric artifact (Zhou et al., 2010).

produced, and for that reason no such effective and autonomous causal theory has been proposed.

Flynn Effect researchers must observe the holistic condition of their field and come to terms with the likely conclusion: the Flynn Effect causal debate cannot be solved while researchers remain in causal theory factions. Instead of focusing on diminutive weaknesses of rival theories, investigators should team together and propose multifaceted theories in order to more accurately explain the complexities of the Effect. The implications of the Flynn Effect on modern society are large, so researchers' opposition to alternative methods of causal investigation would be unwise. Rather, the pressing importance of the Flynn Effect on special education and high-stakes cases should motivate researchers to collaborate and inspire the construction of the most compelling and comprehensive causal theory yet.

Flynn Effect Implications

Although there is still much to learn about the causes of the Flynn Effect, many spheres of public policy have been affected by the implications of the accepted aggregate increase of IQ scores as described by the Effect. Flynn's (1984) original findings have perhaps the greatest consequences for those who rely on IQ scores for special education assessment, because the majority of IQ tests in the United States are administered to students to determine their academic placement (Kanaya et al., 2003). Specifically, populations of children who qualify for mental retardation (MR) or learning disability (LD) must do so with an IQ score from the Wechsler or Stanford-Binet intelligence scales, and the difficulty of obtaining a qualifying score varies greatly during the renorming cycle.

Students with, or being tested for, MR or LD diagnoses are substantially impacted by the IQ test renorming cycle triggered by the Flynn Effect (Truscott & Frank, 2001; Kanaya et al.,

2003). The IQ of this population is fluctuating separate of real cognitive ability simply because of the presence of the Flynn Effect within the lower ranks of the IQ distribution (Kanaya et al., 2003). Consequently, MR or LD identification is entirely dependent upon the year in which the IQ test was taken during the test's renorming cycle, with higher IQ scores on older tests and lower scores on newly renormed tests.

Kanaya et al. (2003) empirically demonstrated this phenomenon through the analysis of 10,800 American special education assessments, confirming the affects of the Flynn Effect within the lower intelligence distribution. The large shift in IQ scores throughout an IQ test renorming cycle resulted in a parallel shift in MR diagnoses; the transition from WISC-R to WISC-III resulted in a mean difference of nearly three IQ points, so students tested on the newer version of the test, were more likely to qualify for a learning disability. In fact, the research conducted by Kanaya et al. (2003) found that the number of MR diagnoses nearly tripled after the introduction of a new IQ test norm.

The Individuals with Disabilities Education Act (Public Law 94-142, 1975) mandated that the majority of special education assessments include an IQ examination, with qualifying students retested every three years for recertification (Ceci & Kanaya, 2010). The implications of the Flynn Effect on longitudinal IQ testing become increasingly important in this system: students serially retested on the same norm will yield progressively higher IQs whereas student IQs will drop dramatically if a new norm is introduced during the cycle of reevaluations (Ceci & Kanaya, 2010).

The consequences of special education diagnoses can extend far beyond the classroom, and so does the relationship between MR classifications and the Flynn Effect. The magnitude of the Flynn Effect on national policies is broad (see Kanaya, Scullin, & Ceci, 2003), but arguably

the most high-stakes application of a MR diagnosis is for defendants in capital murder cases as determined by the *Atkins v. Virginia* (2002) Supreme Court ruling. In the course of appeals following a murder case involving Daryl Atkins, the Supreme Court ruled that executing the mentally retarded violates the Eighth Amendment ban on cruel and unusual punishment (Walker, 2008). Atkins presented an archived record of a special education assessment on which he scored 59, qualifying him for mild mental retardation with an IQ score under the established mental illness cutoff at 70 points (Walker, 2008) thus disqualifying him from death penalty eligibility.

Flynn Effect researchers (e.g. Ceci, Scullin, & Kanaya, 2003; Fletcher, Karla, Stuebing, & Hughes, 2010; Hagan, Drogin, & Guilmette, 2010) vehemently responded to the *Atkins v. Virginia* (2002) case, considering the Effect's direct implications on the Court's decision. The precedent of the case has sparked another debate within the field of Flynn Effect research as the practice of retroactive adjustment of IQ scores in comparable high-stakes decisions is considered. One faction of researchers (e.g. Gresham & Reschly, 2011; Reynolds, Niland, Wright, & Rosenn, 2010) insists that IQ scores should be adjusted for the Flynn Effect in high stakes decisions because, without adjustment, MR diagnostic standards change over the course of IQ test renorming cycles (Fletcher et al., 2010). According to this perspective, the Flynn Effect aggregate mean IQ increase of 0.3 points per year is "sufficiently precise" to justify retroactive adjustments in high-stakes cases (Fletcher et al., 2010, p. 472); additionally, widespread acceptance of the Flynn Effect within the field as evidenced by IQ test publishers' recent tendency to regularly renorm IQ tests demonstrates that the Effect is valid and should be applied to legal decisions.

The opposing group of researchers (e.g. Ceci et al., 2003; Hagan et al., 2010) fervently resists any retroactive adjustment of IQ scores, especially in high-stakes cases. Researchers are still far from understanding the exact nature of the Flynn Effect across IQ tests, levels and types of intelligence, and time. Until the understanding of the Effect is more complete, it is not acceptable to subtract an aggregate mean of 0.3 points per year from an individual score in order to determine a more precise assessment of IQ (Ceci & Kanaya, 2010).

The black box of the Flynn Effect is far from illuminated (Ceci & Kanaya, 2010), and until Flynn Effect researchers can come to a clear and complete understanding of the Effect, is application within public policy should be limited. IQ score adjustment in high-stakes decisions may someday be appropriate, but until individual and aggregate responses to the Flynn Effect can be reconciled, retroactive adjustment of IQ scores in high-stakes decisions is ill-advised. Such important decisions cannot rest on "sufficiently precise" data, but instead adjustments for the Flynn Effect should only be applied in the case of exact understanding of the magnitude and nature of the Effect.

Conclusion

Clearly, the social ramifications of the Flynn Effect are extensive, and the concrete application of the Effect should motivate investigators toward the search for a more holistic causal explanation. The current state of the postulation of hypothetical causal theories is not moving Flynn Effect research any closer to true understanding of the nature of the Effect, so investigators must pragmatically accept the necessary shift toward a more collaborative approach. Unfortunately, the Flynn Effect was prematurely adopted within public discussion before it was properly understood empirically, but continued research into the nature and

magnitude of the Effect can have significant implications in the lives of many (Kanaya & Ceci, 2010).

The causal debate surrounding the nature of the Flynn Effect provides an exciting frontier for researchers to explore and conquer. With any luck, alternative methods of investigation and collaborative researchers will eventually unlock the mystery of the Flynn Effect such that its true nature will be understood.

References

Ang, S., Rodgers, J. L., & Wanstrom, L. (2010). The Flynn Effect within subgroups in the U.S.: Gender, race, income, education, and urbanization differences in the NLSY-Children data. *Intelligence, 38*(4), 367-384. doi:10.1016/j.intell.2010.05.004

Atkins v. Virginia, 534 U.S. 1122 (2002).

Barnett, W. S., Brown, K., & Shore, R. (2004). *The universal debate: Should the United States have preschool for all?* National Institute for Early Education Research. Retrieved from http://nieer.org/resources/policybriefs/6.pdf

Beaujean, A. A., & Osterlind, S. J. (2008). Using Item Response Theory to assess the Flynn Effect in the National Longitudinal Study of Youth 79 Children and Young Adults data. *Intelligence, 36*(5), 455-463. doi:10.1016/j.intell.2007.10.004

Billingsley, B. S., & Cross, L. H. (1992). Predictors of commitment, job satisfaction, and intent to stay in teaching: A comparison of general and special education factors. *The Journal of Special Education, 25*(4), 453-471. Retrieved from http://sed.sagepub.com/content/25/4/453.full.

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 neurobiology of the prefrontal cortex. *Intelligence, 33*, 93-106. doi:10.1016/j.intell.2004.07.008

Bolen, L. M., Aichinger, K. S., Hall, C. W., & Webster, R. E. (1995). A comparison of the performance of cognitively disabled children on the WISC-R and WISC-III. *Journal of Clinical Psychiatry, 51*, 89-94. doi:10.1002/1097-4679(199501)51:1<89::AID-JCLP2270510114>3.0.CO;2-L

Brand, C. R., Freshwater, S., & Dockrell, W. B. (1989). Has there been a ‘massive’ rise in IQ levels in the West? Evidence from Scottish children. *Irish Journal of Psychology, 10*(3), 388-393.

Cattell, R. B. (1968). The theory of fluid and crystallized intelligence: Its relationship to culture free tests and its verification in 9-12 yr. old children. *Bollettino di Psicologia Applicata, 88-90*, 3-22.

Ceci, S. J., & Kanaya, T. (2010). 'Apples and oranges are both round': Furthering the discussion on the Flynn effect. *Journal of Psychoeducational Assessment, 28*(5), 441-447. doi:10.1177/0734282910373339

Ceci, S. J., Scullin, M., & Kanaya, T. (2003). The difficulty of basing death penalty eligibility on IQ cutoff scores for mental retardation. *Ethics & Behavior, 13*, 11-17. doi:10.1207/S15327019EB1301_03

Daley, T. C., Whaley, S. E., Sigman, M. D., Espinosa, M. P., & Neumann, C. (2003). IQ on the rise: The Flynn effect in rural Kenyan children. *Psychological Science, 14*(3), 215-219. doi:10.1111/1467-9280.02434

Delgado, C. L. (2006). Animal source foods to improve micronutrient nutrition and human functioning in developed countries. *Journal of Nutrition, 133*(11), 3907-3910. Retrieved from http://jn.nutrition.org/content/133/11/39

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

Eysenck, H. J., & Schoenthaler, S. J. (1997). Raising IQ level by vitamin and mineral supplementation. In R. J. Sternberg, & E. L. Grigorenko (Eds.), *Intelligence, heredity, and environment* (pp. 363-392). New York, NY: Cambridge University Press.

Firestone, W. A., Schorr, R. Y., & Monfils, L. F. (2004). *The ambiguity of teaching to the test standards, assessment, and educational reform*. Mahwah, NJ: L. Erlbaum Associates Publishers.

Fletcher, J. M., Stuebing, K. K., & Hughes, L. C. (2010). IQ scores should be corrected for the Flynn effect in high-stakes decisions. *Journal of Psychoeducational Assessment, 28*(5), 469-473. doi:10.1177/0734282910373341

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

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

Flynn, J. R. (1994). IQ gains over time. In R. J. Sternberg (Ed.), *The encyclopedia of human intelligence* (pp. 617-623). New York: Macmillan.

Flynn, J. R. (1996). What environmental factors affect intelligence: The relevance of IQ gains over time. In D. K. Detterman (Ed.), *The environment* (pp. 17-29). Westport, CT: Ablex Publishing.

Flynn, J. R. (1998). IQ gains over time: Toward finding the causes. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 25-66). Washington, DC: American Psychological Association. doi:10.1037/10270-001

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist, 54,* 5-20. doi:10.1037/0003-066X.54.1.5

Greenfield, P. M. (1998). The cultural evolution of IQ. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 81-123). Washington, DC: American Psychological Association. doi:10.1037/10270-003

Gresham, F. M., & Reschly, D. J. (2011). Standard of practice and Flynn Effect testimony in death penalty cases. *Intellectual and Developmental Disabilities, 49*(3), 131-140. doi:10.1352/1934-9556-49.3.131

Hagan, L. D., Drogin, E. Y., & Guilmette, T. J. (2010). IQ scores should not be adjusted for the Flynn effect in capital punishment cases. *Journal of Psychoeducational Assessment, 28*(5), 474-476. doi:10.1177/0734282910373343

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class in American life*. New York: Free Press.

Ingersoll, R. M. (2002). The teacher shortage: A case of wrong diagnosis and wrong prescription. *NASSP Bulletin, 86*(16), 16-31. Retrieved from http://bul.sagepub.com/content/86/631/16.full.

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

Kanaya, T., & Ceci, S. J. (2011). The Flynn Effect in the WISC subtests among children tested for special education services. *Journal of Psychoeducational Assessment, 29*(2), 125-136. doi:10.1177/0734282910370139

Kanaya, T., Ceci, S. J., & Scullin, M. H. (2003). The rise and fall of IQ in special ed: Historical trends and their implications. *Journal of School Psychology, 41*(6), 453-465. doi:10.1016/j.jsp.2003.08.003

Kanaya, T., Scullin, M. H., & Ceci, S. J. (2003). The Flynn Effect and U.S. policies: The impact of rising IQ scores on American society via mental retardation diagnoses. *American Psychologist, 58*(10), 778-790. doi:10.1037/0003-066X.58.10.778

Kaufman, A. S. (2010). 'In what way are apples and oranges alike?' A critique of Flynn's interpretation of the Flynn Effect. *Journal of Psychoeducational Assessment, 28*(5), 382-398. doi:10.1177/0734282910373346

Kirp, D. L. (2007). *The sandbox investment the preschool movement and kids-first politics*. Cambridge, Mass.: Harvard University Press.

Lynn, R. (1982). IQ in Japan and the United States shows a growing disparity. *Nature, 297*(5863), 222-223. doi:10.1038/297222a0

Lynn, R. (1987). Japan: Land of the rising IQ. A reply to Flynn. *Bulletin of the British Psychological Society, 40*, 464-468.

Lynn, R. (1989). Positive correlation between height, head size, and IQ: A nutrition theory of the secular increases in intelligence. *British Journal of Educational Psychology, 59*(3), 372-377. doi:10.1111/j.2044-8279.1989.tb03112.x

Lynn, R. (1990). The role of nutrition in secular increases in intelligence. *Personality and Individual Differences, 11*(3), 273-285. doi:10.1016/0191-8869(90)90241-I

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. doi:10.1037/10270-006

Mingroni, M. A. (2004). The secular rise in IQ: Giving heterosis a closer look. *Intelligence, 32*, 65-83. doi:10.1016/S0160-2896(03)00058-8

Mingroni, M. A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn Effect and other trends. *Psychological Review, 114*(3), 806-829. doi:10.1037/0033-295X.114.3.806

Mosley, W. H., & Chen, L. C. (1984). An analytical framework for the study of child survival in developing countries. *Population and Development Review, 10*, 25-45. Retrieved from http://www.jstor.org/stable/2807954

Neisser, U. (1998). Introduction: Rising test scores and what they mean. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures* (pp. 3-22). Washington, DC: American Psychological Association.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77-101. doi:10.1037/0003-066X.51.2.77

Nettelbeck, T., & Wilson, C. (2004). The Flynn Effect: Smarter not faster. *Intelligence, 32*, 85-93. doi:10.1016/S0160-2896(0ro3)00060-6

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110 § 115, Stat. 1425 (2002).

Popham, W. J. (2001). Teaching to the test. *Educational Leadership, 58*(6), 16-20.

Popkin, B. M. (2006). Global nutrition dynamics. *The American Journal of Clinical Nutrition, 84*(2), 289-298. Retrieved from http://www.ajcn.org/content/84/2/289.f

Pritchard, M. E., & Wilson, G. S. (2003). Using Emotional and Social Factors to Predict Student Success. *Journal of College Student Development, 44*(1), 18-28. Retrieved from http://wesamuels.org/pdfs/marked_up

Public Law 94-142. (1975). *Education for All Handicapped Children Act*. U.S. Congress.

Reynolds, C. R., Niland, J., Wright, J. E., & Rosenn, M. (2010). Failure to apply the Flynn correction in death penalty litigation: Standard practice of today maybe, but certainly

malpractice of tomorrow. *Journal of Psychoeducational Assessment*, *28*(5), 477-481. doi:10.1177/0734282910373348

Rodgers, J. L. (1999). A critique of the Flynn Effect: Massive IQ gains, methodological artifacts, or both? *Intelligence, 26*, 337-356.

Rodgers, J. L. (2010). The epistemology of mathematical and statistical modeling: A quiet methodological revolution. *American Psychologist, 65*, 1-12.

Rodgers, J. L., & Wanstrom, L. (2007). Identification of a Flynn Effect in the NLSY: Moving from the center to the boundaries. *Intelligence, 35*, 187-196. doi:10.1016/j.intell.2006.06.002

Rosano, A., Botto, L. D., Botting, B., & Mastroiacovo, P. (2000). Infant mortality and congenital anomalies from 1950 to 1994: An international perspective. *Journal of Epidemiology and Community Health*, *54*, 660-666. Retrieved from http://jech.bmj.com/content/54/9/660.sh

Schoenthaler, S. J., Amos, S. P., Eysenck, H. J., Peritz, E., & Yudkin, J. (1991). Controlled trial of vitamin-mineral supplementation: Effects on intelligence and performance. *Personality and Individual Differences, 12*(4), 351-362. doi:10.1016/0191-8869(91)90287-L

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

Steen, R. (2009). *Human intelligence and medical illness: Assessing the Flynn Effect*. New York, NY: Springer Science + Business Media. doi:10.1007/978-1-4419-0092-0

Striffler, S. (2005). *Chicken: The dangerous transformation of America's favorite food*. New Haven: Yale University Press.

Sundet, J. M., Barlaug, D. G., & Torjussen, T. M. (2004). The end of the Flynn Effect? A study of secular trends in mean intelligence scores of Norwegian conscripts during half a century. *Intelligence, 32*(4), 349-362. doi:10.1016/j.intell.2004.06.004

te Nijenhuis, J., & van der Flier, H. (2007). The secular rise in IQs in the Netherlands: Is the Flynn Effect on g? *Personality and Individual Differences, 43*(5), 1259-1265. doi:10.1016/j.paid.2007.03.016

Teasdale, T. W., & Berliner, P. (1991). Kindergarten attendance in relation to educational level and intelligence in adulthood: A geographical analysis. *Scandinavian Journal of Psychology, 32*(4), 336-343. doi:10.1111/j.1467-9450.1991.tb00884.x

Teasdale, T. W., & Owen, D. R. (1987). National secular trends in intelligence: A twenty-year cross-sectional study. *Nature, 325*(700), 119-121. doi:10.1038/325119a0

Teasdale, T. W., & Owen, D. R. (2005). A long-term rise and recent decline in intelligence test performance: The Flynn Effect in reverse. *Personality and Individual Differences, 39*(4), 837-843. doi:10.1016/j.paid.2005.01.029

Truscott, S. D., & Frank, A. J. (2001). Does the Flynn effect affect IQ score of students classified as LD? *Journal of School Psychology, 39*(4), 319-334. doi:10.1016/S0022-4405(01)00071-1

Urdan, T. C., & Paris, S. G. (1994). Teachers' perceptions of achievement tests. *Educational Policy, 8*(2), 137-156. Retrieved from http://epx.sagepub.com/content/8/2/137.f

Walker, T. G. (2008). *Eligible for execution: The story of the Daryl Atkins case*. Washington, D.C.: CQ Press.

Wei, J., & Putallaz, M. (in press). The Flynn Effect puzzle: A 30-year examination from the right tail of the ability distribution provides some missing pieces. *Intelligence* (2011). doi:10.1016/j.intell.2011.07.006

Weiss, L. G. (2010). Considerations on the Flynn Effect. *Journal of Psychoeducational Assessment, 28*(5) 482-493. doi:10.1177/0734282910373572

Wicherts, J. M., Dolan, C. V., Hessen, D., Oosterveld, P., Baal, G. C. M., van Boomsma, D. I., et al. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn Effect. *Intelligence, 32*(5), 509-537. doi:10.1016/j.intell.2004.07.002

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. doi:10.1037/10270-004

Zhou, X., Zhu, J., & Weiss, L. G. (2010). Peeking inside the "Black Box" of the Flynn Effect: Evidence from three Wechsler instruments. *Journal of Psychoeducational Assessment, 28*(5), 399-411. doi:10.1177/0734282910373340