Applied Neuropsychology
2007, Vol. 14, No. 4, 262–266

Commentary: The Flynn Effect Revisited

Elbert W. Russell
Miami, Florida

The Flynn Effect postulates that intelligence is increasing over time. However, as an environment becomes optimal, a plateau occurs when general growth becomes largely determined by genetics. There is evidence that such a plateau is occurring for intelligence in countries with optimal social environments. In the United States, examination of adult Wechsler test scores between normings indicates a reduction of the FSIQ increase such that average FSIQ would plateau about 2024. However, the WAIS-III norming process eliminates many types of subjects with possible brain impairment. This probably raises the average FSIQ level. With an increase of only 1 FSIQ point in 16 years, a plateau in the Flynn Effect would have been reached in 2004.

Key words: Flynn Effect, intelligence, IQ, Wechsler

Since the late 1970s James R. Flynn (1999) has produced a remarkable set of studies, which demonstrated that the intelligence of people in various countries of the world has been increasing over time. This has been called the Flynn Effect. The rate of increase appeared to be about 0.3 FSIQ points a year for the United States (Flynn, 1999, p. 6). As a result IQs on the same intelligence test would be expected to increase with each new generation.

The Impact of the Flynn Effect

Since intelligence testing has now become a major force in our society, any change in measured intelligence will have wide-ranging consequences. In regard to test construction itself there are recommendations that intelligence tests should be renormed at regular intervals. These recommendations include the Standards for Educational and Psychological Testing (AERA et al. 1999, p. 59, 4.18). The authors of the WAIS-III have stated (Wechsler, 1997, pp. 9–10) that measured IQ is increasing at a 0.3 points each year as it had between 1955 (Wechsler, 1955) and 1981 (Wechsler, 1981, p. 47). This makes "periodic updating of the norms essential" (WAIS-III, 1997, p. 9). In accordance, most companies that produce intelligence tests are beginning to do periodic renorming. All of these recommendations for renorming IQ tests at intervals are based on the assumption that a Flynn Effect will continue indefinitely.

Such renorming of intelligence tests, while good for the test companies, has many consequences for a society so heavily dependent upon intelligence tests. It affects education, social security, the death penalty, and military conscription standards (Kanaya, Scullin, & Ceci, 2003). Many decisions made in the educational system are dependent upon IQ measures. The admission of students to special education classes either for superior ability or for the mentally handicapped is dependent upon IQ scores. A child one year might not be eligible for special education class but in the next year with a new renormed test the same child would be eligible although his ability remained the same. The eligibility for mental retardation services with all its consequences is determined by IQ levels (Kanaya et al., 2003).

The consequences for criminal law are even more critical. A person who is mentally retarded is treated differently in legal cases than those with a normal IQ (Kanaya et al., 2003). In some cases a person's life may depend upon the few points difference produced by renorming an intelligence test. A person could be saved from death penalty

Address correspondence to Elbert W. Russell, 9350 S. Dixie Highway, Suite PH3, South Miami, FL 33156. E-mail: ewr@bellsouth.net

because they were "mentally retarded." However, if the psychologist had not yet bought the renormed version of the test the person might be executed.

The effect of renorming, which constitutes changing an intelligence test, has dramatic and dire effects on research. The advent of a renormed Wechsler intelligence test may consign hundreds even thousands of research studies to the wastebasket. This means that every decade or so, possibly a quarter of the psychological studies involving cognition and neuropsychology are scrapped. This is particularly true in regard to research involved in forensic cases. Studies in other fields can say that the results demonstrated for an older version of the Wechsler test also hold for a newer version. This is not true for forensic cases where attorneys demand the most recent version of an intelligence test.

There is one consideration concerning the Flynn Effect that has not been widely discussed. Generally the measured IQ of mature people remains relatively constant during their lifetime (Mackintosh, 1998, pp. 60–62) although there are maturation effects (Kaufman, 2000, 2001). That is, the Flynn Effect is a cohort effect. The IQs of each succeeding generation may increase but the IQ of an individual remains the same relative to his/her cohort. Consequently, the renorming of intelligence tests will produce scales in which age and cohorts effects on ability are confounded. If a person's ability is measured by IQ tests that have been renormed on a new cohort of subjects, the IQ of older people will be underestimated according to the previous IQ test (Kanaya et al., 2003). That is, for an individual their IQs will become less with each renorming of an intelligence test. The new norming will not give an accurate measure of the individuals ability in absolute terms.

As such, subjects should be tested using the tests that were normed for their cohort—the test that was normed when they were at the average age of the population used to norm a test. This is especially important for such legal purposes as determining whether a person is mentally retarded (Kanaya et al., 2003). Certainly, in forensic situations a person's score on a recently normed version of the Wechsler test should not be compared to a score they obtained on the previous version. They should be compared to the test that they were original administered.

The Effect of Genetics Applied to IQ

The cause of the Flynn Effect has been widely debated (Flynn, 1999). It is evident that the level of intelligence is produced by a combination of inheritance and environmental effects (Van Ijzendoorn & Juffer, 2005). The IQ of individuals follows a very definite pattern of change throughout their lifetime determined by a number of factors but particularly the effect of crystallized and fluid intelligence (Kaufman, 2000, 2001), nevertheless, since the inherited basis of humans is essentially fixed (Mackintosh, 1998, pp. 65–97) the Flynn Effect must be due to some aspect of the human environment (Schneider, 2006).

While the cause is unknown, one of the many possibilities is a change in nutrition. This is supported by the fact that the original generation that was tested in Flynn's review of tests grew up just before or during World War II. Most of the countries in Europe and Japan, which were the basis of these studies, had experienced severe famines during the war and certainly social disruption. The studies examined at a later time involved people who had matured after World War II and whose nutritional requirements were much closer to optimal. The United States had experienced a lesser period of hunger during the depression. The subjects for the early studies thus had experienced nutritional deficiencies, sometimes of a severe nature. A number of studies have found correlations between nutrition and intelligence measures (Sundet, Barlaug, & Torjussen, 2004), although some of Flynn's analyses seemed to contradict this nutritional hypothesis (Flynn, 1999).

Another major factor may be the increasing proportion of the population who obtain a higher level of education. The Wechsler IQ is strongly related to educational level. While part of this may be due to the admission of students with greater ability, nevertheless, this effect is strong enough so that studies of the level of IQ related to maturation and aging, need to take education into consideration (Kaufman, 2000, 2001). For instance, the percent of adults with a college education for the WAIS-R (1981) was 14.4, while the percentage for the WAIS-III (1997) was 25.0 (Kaufman, 2001, p. 137). There are still other possible environmental influences that may have contributed to the Flynn Effect such as a more enriched mental environment.

PLATEAU EFFECT

In biology as one proceeds up a mountain the size of the trees for the same species tend to decrease but at their normal level they tend to grow to about the same height (Krebs, 1985, pp. 114–117). This is the well-known phenomena in biology, that when all environmental requirements for growth exist, organisms reach an asymptote in their growth. The full development of an organism is then determined by the organism's genetics. Additional supplies of the needed requirements usually do not continue to increase the size of the organism. For humans the children of immigrants, such as the children of Japanese, are often larger than their parents. However, this increase does not continue generation after generation. Rather it rapidly reaches a plateau. In an optimal environment the growth of the organism plateaus at a point determined by genetics (Schneider, 2006).

If the level of intelligence is in part inherited and as such is governed by this principle, one would expect that as the environment becomes more optimal for brain development the increase in IQ will slow. When the environmental effects become optimal the intelligence of a population will be determined by genetics and the Flynn Effect will plateau (Schneider, 2006).

Evidence from Studies in Denmark and Norway

Assuming that intelligence is genetically established but is affected by environmental conditions (Kaufman, 2000, 2001), as are most biological characteristics, a plateauing effect would be expected in cultures which have optimal living conditions. Such conditions are more probable in countries that have well-established welfare systems. The Scandinavian countries were the first to establish welfare systems that included the entire population (Schneider, 2006). These welfare systems have generally been in operation since WWII so that even the poorest of the present adult generation would have grown up in a more optimal environment. Consequently, one might expect that a plateauing effect on intelligence would express itself initially in these countries.

In support of this contention studies in both Denmark (Teasdale & Owen, 2005) and Norway (Sundet et al., 2004) have found a plateau in intelligence. This plateauing may also be true of Australia (Cocodia, Kim, Shin, Kim, Ee, Wee, & Howard, 2005). Flynn's own data (1999, p. 8) using the Raven's Progressive Matrices demonstrates that in England there began to be a reduction in the rate of increasing ability for the age groups beginning in 1957. This plateauing is obvious since the same intelligence tests, that provided absolute scores over time, were used. These tests were not renormed every decade or so.

WECHSLER TEST RESULTS

Since there appears to be evidence of plateauing in other countries, one might expect the same result to occur in United States. In order to examine this possibility the most adequately normed intelligence tests in the country, the adult Wechsler tests, will be examined.

Recent Plateauing

The Wechsler tests provide a series of tests that have been periodically renormed from before the 1950s to the present. The differences in these norms are generally accepted by experts in the field (Kaufman, 2001, p. 153). The manual'' for both the WAIS-R (1981) and the WAIS-III (1997) compare these tests to the previous Wechsler adult scales. The WAIS-R, Table 17 (p. 47) compares the WAIS and the WAIS-R using the same set of subjects. This comparison is presented in Table 1 of this present paper. The WAIS (1955) and WAIS-R were published 26 years apart. In this comparison the FSIQ for the WAIS was 111.3 while the WAIS-R was 103.8 a difference of 7.5 FSIQ points. Thus, the FSIQs apparently increased at a rate of 0.288 a year, which is approximately the 0.3 points specified in the WAIS-III manual (1997, p. 8) and by Flynn (1999, p. 6).

However these results are somewhat different from the results found by comparing the group of subjects given the WAIS-R and the WAIS-III (Wechsler, 1997, p. 79). In this case the WAIS-R, FSIQ was 105.8, while the WAIS-III, FSIQ was 102.9 for a difference of 2.9 points. The time between the publications of the WAIS-R the WAIS-III was 16 years. This produced an increase of only 0.118 points per year. Thus, as seen in Table 1, this recent increase is considerably slower than was the difference between the WAIS and the WAIS-R.

As seen in Figure 1, if a linear regression line is drawn between these two rates of increase,

Table 1. *Data Comparing the Increase in Wechsler IQ Points Between 1955 and 1997*

Test	Published	FSIQ	Per Year	Year Normal[a]
WAIS	1955	111.3	–	–
WAIS-R	1981	103.8	–	–
Difference	26 Years	7.5 points	.288	–
WAIS-R	1981	105.8	–	–
WAIS-III	1997	102.9	–	–
Difference	16 Years	2.9 points	.181	2025
Decrease[b] of 1 FSIQ point for WAIS-III				
Difference	16 years	1.9 point	0.118	2004

Note. [a]Predicted year that the extrapolated increase in FSIQ will reach zero.
[b]The decrease in FSIQ produced by subtracting an estimated 1 FSIQ IQ point for the 16 years, caused by the subject selection process.

measured at the midpoints between testings, (Reported rate of increase), the extrapolated reduction in the rate of increase will reach zero about 2024. This means that with these data the IQ in United States will plateau in somewhat less than 20 years.

An Additional Factor

However there is another factor to consider. The WAIS-R and the WAIS-III were not normed in the same way (Kaufman, 2001, p. 136). The WAIS-III norming is not a random sample of United States' population. It is a selected sample. The norming procedure excluded many subjects for a rather large number of reasons, such as alcoholic or drug dependence, a period of unconsciousness of 5 min or more, or history of head trauma (Wechsler, 1997, pp. 21–22). Such an extensive selective process of subjects that might have had a reduced IQ would probably have the effect of raising the mean IQ of the sample.

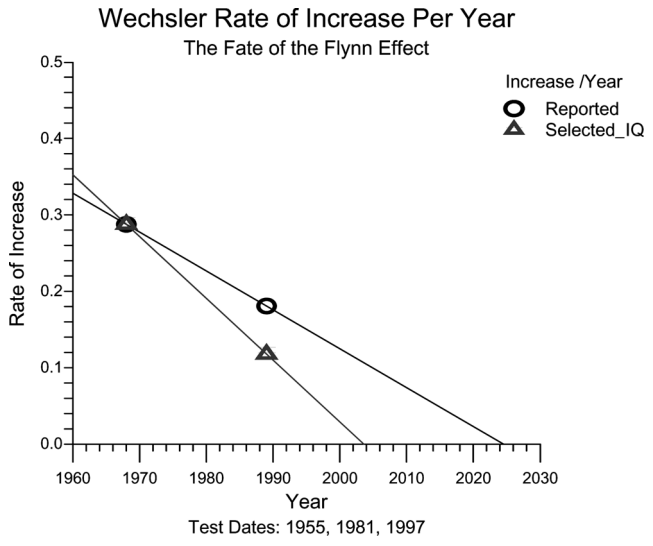

Figure 1. Linear reduction, using mid point measures, in the rate of Wechsler FSIQ increase, between WAIS (1955), WAIS-R (1981), and WAIS-III (1997) compared to rate of reduction if WAIS-III subject selection contributed 1 FSIQ point to FSIQ betwenen WAIS-R (1981) and WAIS-III (1997).

In a recent study Russell (2005) found that all published volunteer samples for the Halstead Reitan Battery norms, that contained an adult Wechsler test published prior to the WAIS-III, had FSIQs that were almost 1 *SD* above average. All of these studies used volunteers and had excluded patients using criteria similar to those employed by the WAIS-III. The previous Wechsler adult scales had not excluded various types of subjects and so represented random samples of the whole population. Since the WAIS-III utilized volunteers that were carefully screened (Wechsler, 1997, p. 22) it does not represent a truly random sample of United States population. Many of the types of subjects that were eliminated would probably have had reduced IQ levels. Thus, there is every reason to suppose that the published intelligence scores of the WAIS-III are significantly higher than if they have been obtained from a random sample.

In regard to the Flynn Effect this means that the actual increase in FSIQ of the United States population between the WAIS-R and the WAIS-III is probably less, even considerably less than that was obtained by the WAIS-III norming. If we assume that the decrease between normings over the entire 16 years, due to this selection procedure was only 1 FSIQ point less than the published FSIQ, the increase in IQ level would have been reduced to 0.118 per year. This is presented in Table 1. If this change in the FSIQ increase is plotted on the same graph, Figure 1, as the

obtained reduction in increase, the extrapolated regression line would place the termination of FSIQ increase at about 2004. That is, the increase in FSIQ would have already plateaued. Thus, it is possible that the Flynn Effect is essentially no longer active in the United States.

CONSEQUENCES

In summary this study presents evidence that the Flynn Effect may not now exist in United States as well as in a number of the other advanced countries. Consequently there is no longer any justification for renorming an intelligence test based on an indefinite increase in intelligence. Intelligence measures would now be invariant overtime and the social consequences of the Flynn Effect would be eliminated.

Of course, tests should be redesigned and normed in order to provide more accurate norms but not because of a Flynn Effect. Also under some conditions, if necessary a test company could provide corrections or adjustments to their general norms for certain segments of a population, whose intelligence might vary from the national average.

REFERENCES

American Educational Research Association, American Psychological Association & the National Council on Measurement in Education (1999). (AERA et al. 1999) *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Cocodia, E. A., Kim, J., Shin, H., Kim, J., Ee, J., Wee, M. S. W., & Howard, R. W. (2005). Evidence that rising population intelligence is impacting formal education. *Personality and Individual Differences*, *35*(4), 797–810.

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist*, *54*, 5–20.

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.

Kaufman, Alan S. (2000). Seven questions about the WAIS-III regarding differences in ability is across the 16 to 89 year life span. *School Psychology Quarterly*, *15*(1), 3–29.

Kaufman, Alan S. (2001). WAIS-III IQs, Horn's theory, and generational changes from young adulthood to old age. *Intelligence*, *29*(2), 131–167.

Krebs, C. J. (1985). *Ecology, the experimental analysis of distribution and abundance* (3rded.) New York Harper Collins.

Mackintosh, N. J. (1998). *IQ and human intelligence*. Oxford, England Oxford University Press.

Russell, E. W. (2005). Norming subjects for the Halstead-Reitan Battery. *Archives of Clinical Neuropsychology*, *20*(4), 479–484.

Schneider, David (2006). Smart as We Can Get? *American Scientist*, *94*, 311–312.

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

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.

Van Ijzendoorn, M. H., & Juffer, F. (2005). Adoption is a successful natural intervention enhancing adopted children's IQ and school performance. *Current Directions in Psychological Science*, *14*(6), 326–330.

Wechsler, D. (1955). *Wechsler adult intelligence scale. Manual*. New York: Psychological Corporation.

Wechsler, D. (1981). *WAIS-R, wechsler adult intelligence scale—revised, manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997). *WAIS-III, WMS-III, technical manual*. San Antonio, TX: Psychological Corporation.