

Secular declines in cognitive test scores: A reversal of the Flynn Effect

Thomas W. Teasdale^{a,*}, David R. Owen^b

^a *Department of Psychology, University of Copenhagen, Øster Farimagsgade 5, 1353 Copenhagen K, Denmark*

^b *Department of Psychology, Brooklyn College, City University of New York, United States*

Received 13 June 2006; received in revised form 30 January 2007; accepted 30 January 2007

Available online 2 March 2007

Abstract

Scores on cognitive tests have been very widely reported to have increased through the decades of the last century, a generational phenomenon termed the ‘Flynn Effect’ since it was most comprehensively documented by James Flynn in the 1980’s. There has, however, been very little evidence concerning any continuity of the effect specifically into the present century. We here report data from a population, namely young adult males in Denmark, showing that whereas there were modest increases between 1988 and 1998 in scores on a battery of four cognitive tests—these constituting a diminishing continuation of a trend documented back to the late 1950’s—scores on all four tests declined between 1998 and 2003/2004. For two of the tests, levels fell to below those of 1988. Across all tests, the decrease in the 5/6 year period corresponds to approximately 1.5 IQ points, very close to the net gain between 1988 and 1998. The declines between 1998 and 2003/4 appeared amongst both men pursuing higher academic education and those not doing so.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Cognitive tests; Secular trend; Flynn Effect

More than 30 years ago Flynn systematically documented evidence, based on test norms of the Stanford–Binet and Wechsler tests, implying that scores on IQ tests had been rising in the United States, through the decades of the last century (Flynn, 1984). Direct evidence came from a review of population studies, in numerous developed countries, where the same test had been used on populations at different times, typically separated by many years, and showing that later generations performed better on tests than earlier generations (Flynn, 1987). The strongest evidence in this category came from countries in which draftees for military service had been tested.

The ‘Flynn Effect’, as it became known, has subsequently been shown to be ubiquitous and there has been much discussion about its causes, with theories ranging from those emphasising social and education changes to those emphasising more biological factors, e.g., health care and nutrition (Neisser, 1998). A recent trend has been the reporting of a Flynn Effect—from direct or indirect evidence—in developing countries (Cocodia et al., 2003; Daley, Whaley, Sigman, Espinosa, & Neumann, 2003; Meisenberg, Lawless, Lambert, & Newton, 2005; Howard, 2005).

Despite the numerous reports, several limitations are also widespread in the relevant literature. A first limitation, as pointed out by Raven (2000), is that many studies involve simply the comparison of mean scores, often using

* Corresponding author.

E-mail address: tom.teasdale@psy.ku.dk (T.W. Teasdale).

scaled and IQ scores such as those calculated in the Wechsler Intelligence tests. This may conceal differential changes at different levels of ability. There is no *prima facie* reason to suppose that the Flynn Effect has involved a uniform shift in mean performance with no change in variance or higher order distributional characteristics such as skewness, and there is some evidence to suggest that gains have not been uniform across all ability levels (Colom, Lluís-Font, & Andres-Pueyo, 2005). A second limitation is that most studies have compared performances on the same tests taken by samples of subjects representative of generations separated by a number of years, typically decades. The limitation of this method is that it is not possible to determine the time course of the effect; a difference between test performance, say, in the 1970's and in the 1990's, could not unambiguously be attributed to a simple linear increase over the two decades. From this follows a third limitation. Even the few studies which have reported data from the present decade, compared to some previous decade, e.g., cannot be taken as evidencing a continuing Flynn Effect up to the present day.

Using data derived from the Danish draft board we have previously reported on substantial gains in test scores, particularly at the low end of the distribution, through the 1960's and 1970's (Teasdale & Owen, 1987, 1989), that by the 1990's they had substantially plateaued (Teasdale & Owen, 2000) and indeed have begun to decline somewhat into the present decade (Teasdale & Owen, 2005). A simultaneous plateau had been reported for Norwegian conscripts by Sundet, Barlaug, and Torjussen (2004) for tests of a range of cognitive abilities. Our own prior report of the present plateau had concerned compound scores on a draft board test summed over four subtests of differing cognitive functions which could be characterised as logical, verbal, numerical and spatial reasoning. The lack of change in the compound score into the present decade could, however, have masked different and possibly opposite trends in these separate abilities. Wicherts et al. (2004) have emphasised the importance of examining Flynn Effect evidence at the level of subtests. The primary objective of the present study has therefore been to examine the recent secular trends in each of these abilities considered separately. A secondary objective has been to explore the relationship between recent changes and educational level.

1. Method

Our data stem from the records of the Danish draft board. There has been conscription in Denmark continuously since the Second World War, and on attaining

the age of 18 or shortly thereafter young men are required to appear before a draft board which assesses their suitability for military service. About 5–10% are exempted from appearing in person, these being largely men who can document a disqualifying illness, e.g., asthma and Scheuermann's disease.

Ever since 1957, and continuing to the present day, the draft board assessment has included an unchanged set of four group-administered tests of cognitive abilities, collectively termed Børge Prien's *Prøve*. The first test, Letter Matrices (19 items, 15 min), resembles Raven's Progressive Matrices with the important difference that patterns of alphabetic letters are used and the correct answer is to be supplied by the testee, rather than chosen from a set of forced-choice alternatives. A Verbal Analogies test (24 items, 5 min) comprises a series of analogies somewhat akin to Miller's Analogies test but where the correct response is to be found in an alphabetically arranged list of 100 words. In a Number Series test (17 items, 15 min), the fifth number following a series of four numbers is to be deduced, and in a Geometric Figures test (18 items, 10 min) a set of complex geometric shapes are to be partitioned into simpler components. All of the tests are scored as the number of correct responses and a total score (0 through 78) is also calculated; this total has been found to correlate 0.8 with the Wechsler Adult Intelligence Scale (Mortensen, Reinisch, & Teasdale, 1989). Further details of the tests are presented elsewhere (Rasch, 1980; Teasdale & Owen, 1989).

Also recorded by the draft board is level of school education. The coding for this was changed substantially in 1991 and the 1988 cohort cannot therefore be compared in this respect with the later two cohorts. The coding is also sensitive to seasonal variation since men tested in the second half of the year are more likely to have completed their ultimate level of schooling. We have therefore here employed a dichotomized index of level of schooling which is independent of season, namely whether or not the subject had attended a 'Gymnasium' (or some equivalent), i.e., an advanced school for 16–18 year-olds leading towards a university entrance and other forms of higher education. Students not attending a Gymnasium typically enter at age 15 or 16 and have shorter and non-academic occupational training courses.

In this report we present data on all men who were tested in 1988 ($n=33,833$), 1998 ($n=25,020$) and in the second half of 2003 together with the first half of 2004, here designated 2003/4 ($n=23,598$) respectively. The difference among the *ns* in these three cohorts is predominantly due to the declining birth-rate in

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات