Forty-year Secular Trends in Cognitive Abilities

T.W. TEASDALE
University of Copenhagen, Copenhagen, Denmark

DAVID R. OWEN
City University of New York, Brooklyn, NY, USA

Changes are shown in the distribution of scores on a set of tests used by the Danish draft board since the late 1950s until the present day. Whereas there were marked gains in earlier decades, especially in the lower end of the distribution, the last 10 years have only seen very modest gains. Such gains as have occurred appear primarily to manifest themselves in a test of visuo-spatial abilities.

It is now over 10 years since Flynn documented, in two highly influential reviews, that performance on intelligence tests had been rising over the preceding decades, both in the United States (Flynn, 1984) and in many other developed countries (Flynn, 1987). The phenomenon was dubbed the "Flynn Effect" by Herrnstein and Murray (1994) by which name it now widely known. The magnitude of the gains found varied from study to study but median figures appeared to fall around 5 IQ points per decade. Discussion of the effect, both with regard to its causes and its implications, has continued unabated to the present date (e.g., Flynn, 1999; Jensen, 1998; Neisser, 1998). There has, however, been a surprising lack of evidence regarding the extent to which the effect has continued through the present decade.

We have earlier reported gains in test performance among young Danish men over a twenty-year period (Teasdale & Owen, 1987) roughly covering the 1960s and 1970s, and thereafter extending this to a 30-year period also including the 1980s (Teasdale & Owen, 1989, 1994). In these studies, we found gains commensurate with those reported for other countries. The objectives of the present study have been first, to examine the extent to which these gains from previous decades have been sustained over the last ten years and second, to examine whether any such gains have been uniform across a range of four different types of cognitive test.

Direct all correspondence to: Dr. T.W. Teasdale, Department of Psychology, University of Copenhagen, Njalsgade 88, DK-2300 Copenhagen S, Denmark. E-mail: teasdale@axp.psl.ku.dk

INTELLIGENCE 28(2): 115–120
ISSN: 0160-2896
Copyright © 2000 by Elsevier Science Inc.
All rights of reproduction in any form reserved.

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

Since 1957 and continuing to the present day, the draft board assessment has included an unchanged set of four group 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 decomposed 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).

In the present report we have combined (anonymous) test data from three cohorts. The first cohort comprises a representative approximately 10 percent sample (n = 71,768) of all Danish men born in the 20 years from 1939 to 1958 (Teasdale & Owen, 1987); over 98 percent of these men were tested between 1957 and 1978. For this sample only the total score, summed across the four subtests, is available. The second and third cohorts comprise all men who were tested in 1988 and in 1998. These numbered 33,833 (born on average in 1969) and 25,020 (born on average in 1979), respectively. The difference between these two numbers is predominantly due to the declining birthrate in Denmark, although in some part it is also due to increasing proportions whose medical unsuitability is registered by the draft board without their need to appear personally. For the 1988 and 1998 cohorts we have obtained scores for each of the four subtests. We have also recorded educational level according to the scale used by the draft board. The scale is broadly based on years of formal education, but it was changed in 1990 with respect to detail, and thus the educational level scales for 1988 and 1998 are not directly comparable.

Results

Fig. 1 shows the mean total score, together with percentiles, as a function of both birthdate (estimated for the cohorts tested in 1988 and 1998) and test date (estimated for the cohort born between 1939 and 1958). It can be seen that the mean gain between test years 1988 and 1998 is diminished in relation to earlier decades; the pattern of means is well described by the quadratic function, estimated mean total score = −14,254.21 + 14.38021

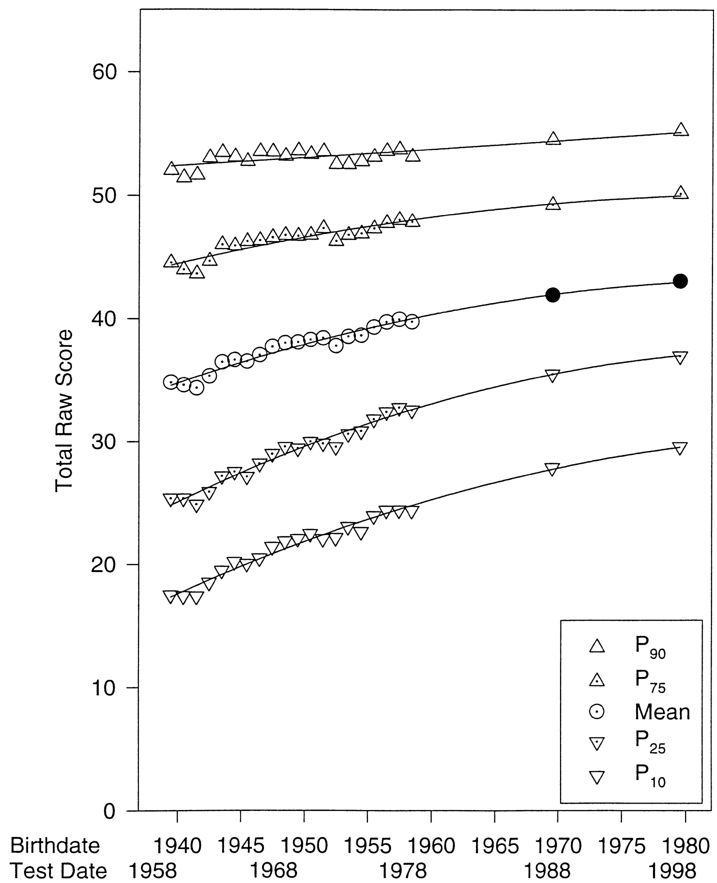

Figure 1. Total test score as a function of birthdate and test date.

× birthdate − 0.00361585 × birthdate^2. Based on this regression equation, the gains in total score across each of the four decades can be expressed as being equivalent to approximately 4, 3, 2, and 1 IQ points. Fig. 1 also shows quadratic regression curves for the 10th, 25th, 75th, and 90th percentiles. The continuation between 1988 and 1998 of a previously reported trend (Teasdale & Owen, 1989) towards relatively greater gains in the lower percentiles, and a very stable 90th percentile, should also be noted.

Prior to our 1988 data collection, there were no complete data on the four separate tests that make up the total scores shown in Fig. 1. Table 1 shows, for the 1988 and 1998 cohorts, summary statistics for the four tests and the total score and Table 2 shows the intercorrelations of the four tests and the total score together with educational level. The stability of the test intercorrelations across the two cohorts is striking. Mean scores have increased for all four tests. However, as shown in Fig. 2, the gain—expressed in terms of 1988 standard deviations—for the Geometric Figures test is more than double that of the

Table 1. Summary Statistics for 1988 and 1998 Tests

Year	*Test*	*M*	*SD*	*Skewness*	*Kurtosis*
1988	Letter Matrices	9.99	2.59	−0.634	0.243
	Verbal Analogies	12.27	4.02	−0.249	0.095
	Number Series	9.61	3.11	−0.560	−0.272
	Geometric Figures	10.06	3.18	0.034	0.166
	Total	41.93	10.44	−0.441	0.134
1998	Letter Matrices	10.18	2.46	−0.671	0.414
	Verbal Analogies	12.53	3.93	−0.350	0.232
	Number Series	9.80	3.04	−0.637	−0.118
	Geometric Figures	10.57	3.18	0.048	0.033
	Total	43.08	10.13	−0.525	0.336

Table 2. Intercorrelations for 1988 and 1998 Tests

Year		*Letter matrices*	*Verbal analogies*	*Number series*	*Geometric figures*	*Total*
1988	Verbal Analogies	0.574				
	Number Series	0.615	0.606			
	Geometric Figures	0.477	0.486	0.453		
	Total	0.798	0.856	0.822	0.745	
	Educational Level	0.407	0.508	0.437	0.307	0.521
1998	Verbal Analogies	0.564				
	Number Series	0.614	0.588			
	Geometric Figures	0.473	0.474	0.433		
	Total	0.794	0.850	0.813	0.743	
	Educational Level	0.462	0.539	0.510	0.319	0.575

other three tests, which differ only slightly among themselves. The Geometric Figures test differs from the other three in several statistical respects. It is the only test not to show a reducing variance between 1988 and 1998, or an increasing negative skew (see Table 1). Furthermore, it has consistently lower correlations with the other three tests and with the total score and educational level (see Table 2).

Discussion

There are two artifactual mechanisms that could explain the slowing rate of gain in total test scores particularly over the last decade, as revealed in Fig. 1. First, it might be a ceiling effect, whereby test scores begin to reach the maximum possible, thus leaving no room for further improvement. There is, however, no strong evidence for this in our data: in none of the four tests was the maximum possible score attained by more than 2 percent of those tested and fewer than 10 percent gave correct solutions to over 80 percent of items. This would, however, still leave the possibility of "glass" ceilings (i.e., below the maximum possible score, created by, for instance, extremely difficult test items). Second, it might be a creeping obsolescence whereby test items become less familiar and intelligible to

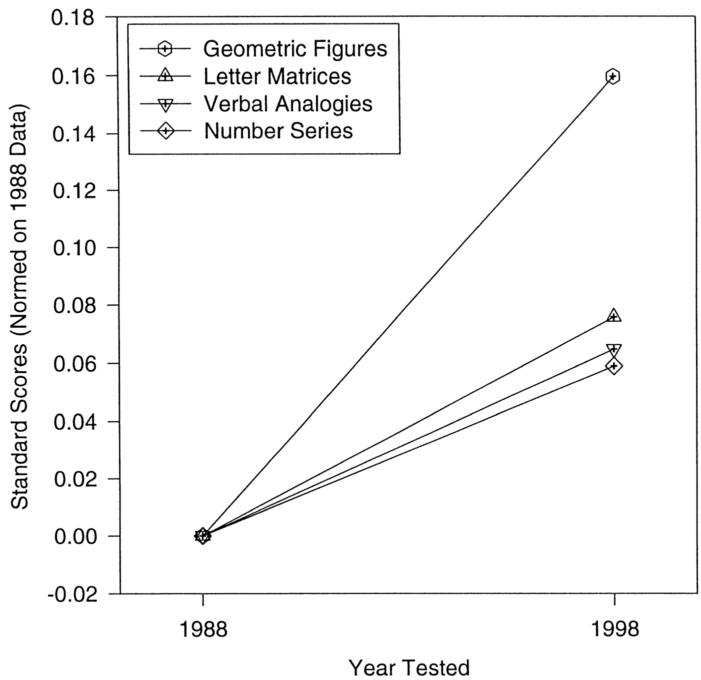

Figure 2. Relative test score gains between 1988 and 1998.

successive generations thereby counteracting any tendency towards further gains. The abstract and "fluid" nature of the four tests makes such argument less plausible in the present instance, and has indeed underlain the continuing use of the tests, unaltered for over four decades. Furthermore, although the change in scales for educational level make comparisons hazardous, the increased correlations between the tests and educational level from 1988 to 1998, shown in Table 2, do not argue for either ceiling effects or obsolescence.

If the slowing rate of gain is not primarily to be explained by these artifactual mechanisms, then an explanation must be sought in a diminution of the factors which have been responsible for the gains themselves. The nature of these factors has been the subject of considerable speculation. Both Jensen (1998) and Neisser (1998) have recently pointed out that hypothesized mechanisms can be grouped as being broadly "biological" or "social." Proposed contributing biological mechanisms have included improved nutrition (Lynn, 1998) and health care as well as outbreeding and heterosis (Jensen, 1998): social mechanisms have included lower family size and thereby birth order (Zajonc & Mullally, 1997) and increasing test sophistication (Brand, 1987). Among social mechanisms, we have emphasized improved education (Teasdale & Owen, 1987, 1989, 1994) including preschool education (Teasdale & Berliner, 1991), a factor that we consider to be prima facie supported by the consistently high correlations found between test scores and educational level, including those found here, and by the fact that the gains were manifested primarily at the lowest levels, where educational reforms of the earlier decades had their greatest impact.

To all of these should be added a consideration of the potential role of motivation, especially where score gains are reported, as here, from draft board tests. (In such cases, reports are limited further, with one notable exception [Flynn, 1998], to males only.) In general, most researchers have leaned towards the view that a variety of factors have been at work, to different degrees and at different times. In the present context, it is tempting to speculate that, as has been suggested by Greenfield (1998), it has been growing exposure to video games and geometrically configured computer screens via operating systems, applications programs and the Internet, that have particularly accelerated the development of visuo-spatial abilities during the last decade.

Acknowledgments: This work was supported by the Danish Research Council for the Humanities and the Inge Lehmann Legacy. It was completed while the first author was at the Medical Research Council's Cognition and Brain Sciences Unit, Cambridge, UK. We thank Stig Meincke, Monica Priestley, Marie Søby, and Vivi Stets for invaluable technical help.

References

Brand, C. R. (1987). Keeping up with the times. *Nature, 328*, 761.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*, 29–51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin, 101*, 171–191.

Flynn, J. R. (1998). Israeli military IQ tests: Gender differences small; IQ gains large. *Journal of Biosocial Science, 30*, 541–553.

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

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

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

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

Lynn, R. (1998). In support of the nutrition theory. In U. Neisser (Ed.), *The rising curve: Long-term gains in IQ and related measures*. Washington, DC: American Psychological Association.

Mortensen, E. L., Reinisch, J., & Teasdale, T. W. (1989). Intelligence as measured by the WAIS and a military draft board group test. *Scandinavian Journal of Psychology, 30*, 315–318.

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*. Washington, DC: American Psychological Association.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Teasdale, T. W., & Berliner, P. (1991). Experience of kindergartens in relation to educational level and cognitive abilities in adulthood: A geographical analysis. *Scandinavian Journal of Psychology, 32*, 336–343.

Teasdale, T. W., & Owen, D. R. (1987). National secular trends in intelligence and education: A 20-year cross sectional study. *Nature, 325*, 119–121.

Teasdale, T. W., & Owen, D. R. (1989). Continuing secular increases in intelligence and a stable prevalence of high intelligence levels. *Intelligence, 13*, 255–262.

Teasdale, T. W., & Owen, D. R. (1994). 30 year secular trends in the cognitive abilities of Danish male school leavers at a high educational level. *Scandinavian Journal of Psychology, 35*, 328–335.

Zajonc, R. B., & Mullally, P. R. (1997). Birth order: Reconciling conflicting effects. *American Psychologist, 52*, 685–699.