ARTICLE IN PRESS

Intelligence xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Intelligence

journal homepage: www.elsevier.com/locate/intell

IQ decline and Piaget: Does the rot start at the top?

James R. Flynn*, Michael Shayer

University of Otago, New Zealand

A B S T R A C T

The IQ gains of the 20th century have faltered. Losses in Nordic nations after 1995 average at 6.85 IQ points when projected over thirty years. On Piagetian tests, Britain shows decimation among high scorers on three tests and overall losses on one. The US sustained its historic gain (0.3 points per year) through 2014. The Netherlands shows no change in preschoolers, mild losses at high school, and possible gains by adults. Australia and France offer weak evidence of losses at school and by adults respectively. German speakers show verbal gains and spatial losses among adults. South Korea, a latecomer to industrialization, is gaining at twice the historic US rate.

When a later cohort is compared to an earlier cohort, IQ trends vary dramatically by age. Piagetian trends indicate that a decimation of top scores may be accompanied by gains in cognitive ability below the median. They also reveal the existence of factors that have an atypical impact at high levels of cognitive competence. Scandinavian data from conventional tests confirm the decimation of top scorers but not factors of atypical impact. Piagetian tests may be more sensitive to detecting this phenomenon.

1. Introduction

Scandinavian data suggest that there has been an IQ decline in some advanced nations and that it began about 1995. The US is of interest because it appears to the leading exception to IQ decline. Great Britain may be more important because it allows us to compare trends on Raven's Progressive Matrices with those on Piagetian tests. We will compare trends in all of these nations to trends in advanced nations elsewhere.

Our comparative analysis does not included developing nations because: either the samples are local and the trends best explained by factors peculiar to the recent history of the locale (Brazil); or the samples are atypical (Saudi Arabia).

For completeness's sake, Colom, Flores-Mendoza, and Abad (2007) compared children aged 7 to 11 who took the Draw-a-Man test in the city of Belo Horizonte, Brazil: 499 in 1930 and 710 in 2002. Over the 72 years, they gained 17 IQ points for a rate of 0.236 points per year. Bandeira, Costa, and Arteche (2012) compared children aged 6 to 12 from the state schools of Porto Alegre. They compared 294 assessed during 1980 decade with 203 assessed during the 2000 decade on Draw-a-Man and found no gain. They also compared 562 children assessed during the 1990 decade and 243 assessed during the 2000 decade on Raven's Coloured Matrices and found the same. The authors believe that the most likely cause of the trend is that the management of primary state schools is now the responsibility of local councils, rather than of the state education department. As a result, primary schools suffered a decline in their budget and perhaps worse overall quality of primary education. In Saudi Arabia, Bakhiet, Baraket, and Lynn (2014) found substantial gains by schoolchildren from 1999 to 2013 on the Coloured Matrices that matched those of Colom et al. (2007) from 1977 to 2010 on the Standard Progressive Matrices. But Bakhiet's samples were boys who were deaf.

After our analysis, we will suggest two tentative hypotheses. First, trends on conventional tests show those at most risk of IQ decline are high school students aged 14 to 18. However, Piagetian results in Britain imply losses at earlier ages. Second, Piagetian tests signal something extra: conflicting trends between top scorers (those at the highest or formal level of cognitive development) and those in the early stages of the next level (concrete generalization). Large losses at the formal level may be accompanied by gains at the concrete level. We will argue that conventional IQ tests can show this phenomenon but are less likely to do so.

2. Approach to assigning causes

Our approach will be essentially sociological, that is, matching IQ trends with changes in the cognitive demands of social institutions like the family, the schools and the world of work. Over several generations, there is no doubt that dysgenic selection lowered the quality of genes for intelligence and increased migration had an adverse effect. Kong et al. (2017) found that genes predictive of educational attainment and *g* have been declining steadily across birth cohorts in Iceland. Woodley

* Corresponding author.
E-mail addresses: jim.flynn@otago.ac.nz (J.R. Flynn), m.shayer@btinternet.com (M. Shayer).

https://doi.org/10.1016/j.intell.2017.11.010
Received 22 September 2017; Received in revised form 20 November 2017; Accepted 27 November 2017

Table 1
Mental test trends in Scandinavia divided into recent gains and losses (measured in IQ points)[a].

Nation	Test	Years compared	IQ gain per year	IQ loss per year	Total loss	Projected 30 yrs.
Finland	Word	1988–1997	0.15	–		
Finland	Word	1997–2009	–	− 0.39	4.66	11.65
Finland	Number	1988–1997	0.37	–		
Finland	Number	1997–2009	–	− 0.27	3.25	8.13
Finland	Shapes	1988–1997	0.76	–		
Finland	Shapes	1997–2009	–	− 0.09	1.07	2.68
Average			**0.43**	**− 0.25**		**7.49**
Denmark	Verbal analogies	1988–1998	0.10	–		
Denmark	Verbal analogies	1998–2003/4	–	− 0.22	1.22	6.65
Denmark	Number series	1988–1998	0.09	–		
Denmark	Number series	1998–2003/4	–	− 0.32	1.77	9.65
Denmark	Geometrical figures	1988–1998	0.24	–		
Denmark	Geometrical figures	1998–2003/4	–	− 0.20	1.10	6.00
Denmark	Letter matrices	1988–1998	0.11	–		
Denmark	Letter matrices	1998–2003/4	–	− 0.12	0.66	3.60
Average			**0.135**	**− 0.215**		**6.48**
Norway	Word similarities	1977–1993	0.01	–		
Norway	Word similarities	1993–2002		− 0.27	2.46	8.19
Norway	Arithmetic	1977–1993	0.11	–		
Norway	Arithmetic	1993–2002		− 0.41	3.73	12.43
Norway	Figures	1977–1993	0.39	–		
Norway	Figures	1993–2002		(+ 0.03)	(+ 0.24)	(+ 0.80)
Average			**0.17**	**− 0.22**		**6.50**
Sweden	Instructions	1982/83–1992/93	0.18	–		–
Sweden	Synonyms	1982/83–1992/93	(− 0.06)	–		–
Sweden	Metal folding	1982/83–1992/93	0.23	–		–
Average			**0.12**	–		–
Overall Average			**0.21**	**− 0.23**		**6.85**
Sweden	Technical comp.	1982/83–1992/93	0.01	–		

[a] We have calculated using the means and SDs presented in tables and in some cases, our values in terms of IQ points differ from those reported.

of Menie, Peñaherrera, Fernandes, and Figueredo (2017) did a synthetic literature search that yielded a total of 66 observations of secular IQ decline from 13 countries with a combined sample size of 302,234, which spanned the 87 years from 1920.5 to 2007.5. They found that high levels of immigration were correlated with IQ decline, particularly when the decline was measured by tests with a high g loading.

However, our analysis focuses on a relatively brief period from about 1980 to 2014, or about one generation. Clearly during most of the 20th century social factors were so potent that they overwhelmed other negative trends, hence IQ gains over time.

There is every reason to believe that the same thing occurs in the transitional generation we analyze, except that the social factors may now have turned from positive to negative. It is true that at the end we speak of prospects over the 21st century in sociological terms but only because social factors look predictable while dysgenic selection concerning the future is harder to anticipate and measure.

3. The Scandinavia data

Dutton and Lynn (2013) conclude that immigration in Scandinavia during this period is too small to be an important factor. Table 1 traces trends on military tests of 18–19 year olds in Finland, Denmark, Norway, and Sweden (Dutton & Lynn, 2013; Rönnlund, Carlstedt, Blomstedt, Lars-Göran Nilsson, & Lars Weinehall, 2013; Sundet, Barlaug, & Torjussen, 2004; Teasdale & Owen, 2008). Between nations, the results are remarkably uniform. The averages by nation in bold make this clear.

The year in which gains turn into losses cluster around 1995. The last rates of gain average at 0.21 IQ points per year over the previous decade (with Norway the period is 16 years). We have included Sweden, which has no recent data but which shows signs of lapsing into losses in 1993. The rates of loss are almost identical, just above or below - 0.23 points/year.

The total losses over the time the data cover look small. But when projected over a generation (30 years), Finland would lose 7.49 points of overall IQ, Denmark 6.48 points, and Norway 6.50 points for an average of 6.85 points. Nonetheless, it would be useful to see if father-son comparisons differ from the results as reported. Table 1 also reports results from the Swedish Technical Comprehension Test. We omitted it from our calculations because its content is atypical of most IQ tests. It poses specialized technical and physical problems.

4. Recent trends the United States

America appears to show a steady rate of gain from 1989 to 1995 to 2002 to 2007 to 2014 on Wechsler tests, which is just about its historic rate of 0.30 IQ points per year. The estimates are often based on small samples that took both an earlier and later edition of a Wechsler test in counterbalanced order. They scored higher against the older norms – thus signaling that the performance of the standardization samples had improved over time.

This method poses the problem of allowing for changed content between tests. Eventually, the Wechsler architects did what needed to be done. Weiss, Gregoire, and Zhu (2016) compared the standardization sample performance on the WISC-IV (normed in 2002.25) with an appropriate sample also tested on the WISC-IV some years later (2013.83). They selected 126 subjects, representative of the WISC-V sample and called the “new” sample. To get 126 controls from the earlier WISC-IV standardization sample, they selected 100 samples of matched subjects, calculated the Full Scale IQ of each and took the median score of these as the IQ of the “old” sample. When the new sample took the WISC-IV, their scores showed a gain of 0.313 points per year. This both adds reliability to other estimates and shows that the US did not suffer losses up to 2014. There are also a number of comparisons between different forms of the Wechsler, between the Wechsler and the Stanford-Binet, and between the SB-4 and SB-5, which confirm these

Table 2
Recent IQ trends in eight nations.

Gains

Nation	Ages	Test(s)	Years compared	IQ pts. gained	Rate per year	Sample	Reference
USA	1–16	WISC-III & WISC-IV	1989–2002	4.63	0.363	244 took both tests	Flynn, 2012a, 238–239
USA	1–16	WISC-IV & WISC - IV	2002.25–2013.83	3.62	0.313	126 matched subjects	Weiss et al., 2016
USA	17–89	WAIS-III & WAIS-IV	1995–2007[a]	3.37	0.281	240 took both tests	Flynn, 2009, Table 2
United Kingdom	5–15	Raven's	1980–2008	6.23	0.221	Analysis of CPM and SPM	Flynn, 2012a, 322
South Korea	5–16	Wechsler tests & K-ABC	1986–1999 (ave. gap = 8.75 years)	6.71	0.767	Small groups took two tests	te Nijenhuis et al., 2012
Netherlands	Adults, children	Various	1971–2005	See Table 3	See Table 3	Various	See Table 3
German speakers	Median Age 42	Vocabulary	1997–2007	3.50	0.350	Meta-analysis 500 + studies	Pietschnig, Voracek, & Formann, 2010
German speakers	6–16	WISC-III & WISC-IV	1997–2006	3.81	0.420	100 took both tests	Petermann & Petermann, 2008

Losses

Nation	Ages	Test(s)	Years compared	IQ pts. lost	Rate per year	Sample	Reference
German speakers	13–42	Spatial	1995–2014	− 9.12	− 0.480	Meta-analysis 96 samples	Pietschnig & Gittler, 2015
France	30–63	WAIS-III & WAIS IV	1999–2008.5	− 3.80	− 0.400	79 took both tests	Dutton & Lynn, 2015
Australia	6–11.92	Raven's Coloured Matrices	1975–2003	− 1.32	− 0.047	693 & 618 from Victoria	Cotton et al., 2005
Estonia	15–16	Raven's	2001–2012	− 8.68? See Table 4	− 0.723? See Table 4	552, 411, 304 from 3 schools	Korgesaar, 2013

[a] The reference puts the WAIS-IV norming date at 2006; 2007 is correct.

results (Flynn, 2012a, pp. 238–239).

Weiss et al. (2016, Table 2) also allows a more detailed analysis of US trends between 2002 and 2014. The gains on Matrix Reasoning (closest to Raven's) were the largest at 5.25 IQ points, a huge gain at the rate of 1.35 points per year, or 4.3 times the size of Full Scale IQ. However, the gains on Working Memory were only 2.10 points. This is a blend of the Arithmetic and Digit Span subtests where gains were 1.65 points and 0.90 points respectively. Digit Span in turn is a blend of Digit Span Forward and Digit Span Backward. The former is mere rote memory: repeating from memory the sequence of digits read out at random. The latte is true working memory: your mind must alter the sequence so you can repeat them in reverse order.

Woodley of Menie and Fernandes (2015) analyzed American data covering 1923 to 2008. They verified that substantial losses on Digit Span Backward had been offset by gains on Digit Span Forward. Thus if we had a true measure of Working Memory (omitting Digit Span Forward), and if the pattern extended from 2008 to 2014, Americans might show both large gains on Raven's and nil or negative gains on Working Memory. Harrison, Shipstead, and Engle (2014) summarize the evidence that there is a strong correlation between the two. But if the correlation reveals a casual or functional relationship between them, then score trends on both show be much the same. Is this a case of correlation deceiving us about causality?

5. Scandinavia and the US

Why has the US not followed the Scandinavian pattern? A more fundamental question is whether trends for these two are inconsistent. The Scandinavian data are all for 18–19 year olds, too young for the world of work, too old to represent school children. Dutch data will soon convince us that trends, even in one nation at one time, can vary by age, depending on whether there is enhancement or deterioration in parenting, schooling, job demands, and the well-being of the aged. If we had WAIS data from Scandinavia, its adults might show gains. If we had WISC data from Scandinavia, it schoolchildren might show gains. In Table 7, we will show that Nordic losses at age 18 do not imply any losses before age 16.

Assuming Nordic and US trends really differ, what social trends *might* foreshadow a general decline in Scandinavia prior to a decline in America. In Scandinavia, the factors that have caused IQ gains may have exhausted their potency. Their educational system is more advanced and may show that at a certain point, schooling has reached a limit in terms of producing more graduates that can generalize and use logic on the hypothetical (mental abilities that pay dividends on IQ tests – Flynn, 2009). Further, good schooling has reached all classes and a more developed welfare state blunts the edge of class. Creation of cognitively demanding administrative jobs may have reached a limit beyond which economic efficiency forbids more "feather-bedding".

What about the world of work as an important source of cognitive development for adults? That is, what have automation and other trends done to the ratio between cognitively demanding jobs and less demanding service work (plus no jobs at all)? As usual economists disagree, but Richard Florida (2014) has provided fascinating US data. From 1990 to 2012: the "working class" share of jobs has steadily declined from 31.4% to 20.5%; the "creative class" share has been pretty stable, rising from 29.3 to 32.0; the "service class" has risen sharply from 39.3 to 48.5. He calls the latter "low-skill jobs", seeing increases in occupations like nursing assistants, personal care aids (whether at home or in nursing homes), retail sales, and food prep workers (McDonalds). Note that the ratio between low skill and creative has risen in favor of the former: 1.35 to one in 1990; and 1.57 to one in 2012.

We will chance our arm and predict that the US ratios measure a trend that will substantially affect all advanced nations. If so, that does not bode well for IQ gains over the current century. Even if unfavorable ratios do bite in all advanced nations, they might vary from one to another. If economic historians trace them for all those we examine, this could shed light on different IQ trends between our eight nations, at least among adults.

In passing, we seen no conflict between the US data and Earl Hunt's analysis in his *Will we be smart enough* (1995). He is quite right that there will be a greater need for people (particularly in the communications sector) that have a sophisticated grasp of new technology and a greater need for people who are flexible enough to adapt as one job is swept out from under them and they must shift to another. But this does not necessarily offset lower cognitive demands from other jobs created. Moreover, he doubts the schools will educate enough people to meet these new needs. And, of course he did not have post-1990 data.

6. Recent IQ trends in eight nations

Table 2 summarizes both the US gains and those of other developed nations. We will analyze these others in turn.

6.1. The United Kingdom

Raven's data cover all the years from 1980 to 2008. The length of this period and the small rate of gain (0.221 points per year) leave open the possibility that at some year the late effects of industrialization began to turn gains into losses, perhaps in 1995 replicating the Scandinavian pattern. Moreover, ages 14–16 are an exception to other ages and show an IQ loss even during the period as a whole (Flynn, 2012a, p. 322). We will soon analyze the UK in the light of contrary trends on Piagetian tests.

6.2. South Korea

te Nijenhuis, Cho, Murphy, and Lee (2012) applies the standard method of using small samples that took two IQ tests normed in different years between 1986 and 1999. These included Kaufman's ABC (Assessment Battery for Children), the WIPPSI, the WISC, and the WAIS (average interval 8.75 years). South Korea is a society that began to industrialize much later than Western European nations and it has done so at a far faster pace than they. If IQ gains begin with the advent of industrialization and end when industrialization becomes counterproductive (in terms of schooling and job creation), Korea would go through the period of IQ gains at a faster rate than Western Europeans. Note that its recent rate of gain stands at 0.723 points per year, a rate roughly double that traditional in the West.

6.3. The Netherlands

The Netherlands is often cited as a nation that began to show IQ losses beginning as early as 1975. However, an analysis of all available data shows that this is suspect. Table 3 is divided into adults, high school students, and preschoolers. The data for all 18-year-old males is a saturation military sample and beyond challenge. In addition, the scores of a random sample of the 1982 recruits were compared to the scores of their own fathers, which confirmed the estimate. The period of gains on Raven's ended in 1981.5, but two straws in the wind indicate that adult Dutch males may have had robust gains until almost the end of the century.

First, the conventional method was used to measure gains from the standardization sample of the WAIS in Holland in 1967.5 to the standardization sample of the WAIS-III in 1998.5 (77 subjects took both tests). We have allowed for the fact that the reported gain was probably two points high because the WAIS-III sample was slightly elite (15.5 points - 2 points = 13.5 points). The rate of 4.35 points per decade is somewhat above the international average. It is about half of the Raven's gain but this is typical: Raven's gains (fluid intelligence) tend to be double gains on crystallized intelligence (Wechsler tests). For what it is worth, applicants to be bus drivers had lesser gains (0.84 points per decade) prior to 1984 but greater gains (6.15 per decade) from 1984 to 1990.

Fragmentary data for other ages are different. High school students of 14 and 16 show a loss of - 3.60 points per decade from 1975 to 1985, which diminishes to - 2.69 points from 1985 to 1995, which virtually disappears during a period extending from 1985 to 2005 (− 0.74 per decade). Indeed the middle value is probably inflated because the sample of 14-olds was substandard — they did not include schools in the larger towns like Amsterdam. Preschoolers, just before their 5th birthday to before their 6th, show a slight gain (0.74 points per decade) from 1981.5 to 1991.5. The later sample was a bit elite in terms of SES. A judicious estimate would be that there was neither gain nor loss. For computations and quality of samples see Box 1.

Are the high school subjects the wave of the future and foreshadow a mild decline, as they become adults? This would be a false inference. Recall that the magnitude of IQ gains/losses varies by age. Scottish data show that when two cohorts (separated by 15 years) were compared at age 11, the latter had gained 3.7 IQ points. When the same cohorts were compared at age 77, the latter had gained 16.5 IQ points (Staff, Hogan, & Whalley, 2014). Flynn (2016a, p. 91) has evidenced that those aged 17.5 to 24 tend to gain at about double the rate of schoolchildren.

Taking the Dutch data at face value may show that families are furnishing a static cognitive environment for preschoolers, high schools are in mild decline, jobs are significantly more cognitively demanding, and the Dutch (like everyone else) have greatly enhanced the quality of life of the aged (better health and more exercise). No data available evidence that the Netherlands has entered a period of IQ decline. We await a really good study that updates trends on military mental tests.

6.4. German speakers (Germany, Austria, and part of Switzerland)

For Vocabulary, Pietschnig et al. (2010) did a meta-analysis of over 500 studies (no. = 45,000 adult individuals) and found a gain from 1971 to 2007 of 0.35 points per year. The rate was uniform thus yielding a 3.5-point gain from 1997 to 2007. For spatial ability,

Table 3
Recent IQ trends in the Netherlands.

Ages	Test	Years compared	IQ points gained or lost	Rate per decade	Sample and no.	Reference
Adults						
18	Raven's	1972–1981.5	+ 8.67	+ 9.13	All males	Flynn, 1987, Table 1
Mean age 40.30	WAIS & WAIS-III	1967.5–1998.5	+ 13.50	+ 4.35	77 took both tests	Wicherts et al., 2004
Bus drivers	GATB (8)	1975.5–1984	+ 0.74	+ 0.87	Applicants: 130 & 1091	Woodley of Menie & Meisenberg, 2013
	GATB (8)	1984–1990	+ 3.69	+ 6.15	Applicants: 1091 & 212	Woodley of Menie & Meisenberg, 2013
High school						
16	GATB	1975–1985	− 3.60	− 3.60	130 & 270	Woodley of Menie & Meisenberg, 2013
	GATB	1985–2005	− 1.49	− 0.74	270 & 498	Woodley of Menie & Meisenberg, 2013
14	DAT	1984–1994.5	− 2.82 (+)	− 2.69 (+)	857 scored vs. DAT norms	Woodley of Menie & Meisenberg, 2013
Preschool						
4.92–5.92	RAKIT	1981.5–1992.5	+ 0.81 (−)	+ 0.74 (−)	208 scored vs. RAKIT norms	Woodley of Menie & Meisenberg, 2013

(+) Next to a loss means the loss should probably be diminished. (−) Next to a gain means it was probably a bit lower than the estimate. See Box 1 for method of computation and description of samples.

Box 1
Netherlands: computations and descriptions of samples.

(1) Wherever "Woodley of Menie & Meisenberg, 2013" appear in Table 3, their Table 1 is the basis of computations. When possible we have divided results for a long time period into two shorter periods, for example, their 1975/76–1988/92 becomes our 1975.5–1984 and 1984–1990.
(2) The gains or losses have been computed by averaging results for all subtests given. The bus drivers took 8 subtests of the GATB (General Ability Test-B): Vocabulary, Space, Arithmetic reasoning, Computation, Tool matching, Form matching, Name comparison, and Mark making. We averaged the gains/losses on these to get an average gain of + 0.058 SDs per decade for 1975.5–1984 and one of + 0.410 SDs per decade for 1984–1990. Take these values times 15 and you get + 0.87 and + 6.15 IQ points, the values presented in Table 3. The method provides a slight underestimate in that the subtests are not perfectly correlated with one another. The same procedure was followed for 16-year olds who took the first seven of the GAT-B subtests named.
(3) For sample quality of the 16-year olds, see te Nijenhuis (2013). For the 14-year olds and preschoolers, see Wicherts et al. (2004). The latter took six subtests from the RAKIT (Revision of the Amsterdam Kinder Intelligence Test): Exclusion, Discs, Hidden figures, Verbal meaning, Learning names, and Idea production.
(4) The WAIS and WAIS-III results are reported and analyzed in Wicherts et al. (2004). The Raven's results are reported and analyzed in Flynn, 1987.

Pietschnig and Gittler (2015) used the 3DC (or three dimensional cubes) test: participants had to decide whether one stimulus cube matches one of six target cubes or that no target cube is correct. A meta-analysis of 96 studies (No. = 13,172; mean ages 13–42) found a loss of 0.48 points per year from 1977 to 2014. They do not give a decade-by-decade break down; but their graph shows a rise to 1995 and a loss thereafter, thus implying that the rate of loss was greater than 0.48 per year between 1996 and 2014. We can find no plausible explanation of why vocabulary and spatial ability trends diverge.

Petermann and Petermann (2008) use the standard method of a small sample (no. = 100) who took both the German WISC-III (normed in 1997) and the German WISC-IV (2006) and got a Full Scale IQ 3.81 points lower on the latter — implying that the performance of standardization samples had risen over time at the familiar rate of 0.42 points per year. From the Vocabulary and Spatial results described above, you would expect Verbal gains to be sizable and Performance gains to be nil or slight. In fact, the Verbal gains were 4.88 points and the Performance gains 4.00 points, both robust. These results might seem to tip Germany toward the column of overall gains but note that the WISC results are for children while the Vocabulary/Spatial results are for adults. As we have seen, adults and children can diverge even at the same time for the same nation.

6.5. France

Dutton and Lynn (2015) use the standard method of a small sample (no. = 79) who took both the French WAIS-III (normed in 1999) and the French WAIS-IV (2008–2009) and got a Full Scale IQ 3.80 points higher on the latter — implying that the performance of standardization samples had fallen over time. There is nothing the matter with this study but we will need other studies to be confident about its results, such as the large number we have in the US. However, in the light of Scandinavian trends, a reversal date in the 1990s cannot be dismissed.

6.6. Australia

Cotton et al. (2005) provide the only study. Raven's Coloured Progressive Matrices was normed in the state of Victoria on children aged from 6 to 11 in both 1975 and 2003. The second sample (618) scored about 1.32 IQ points below the first (693) over the 28 years. The authors stress that the samples are not strictly comparable: the first sampled only state schools but matched census variables; the second sampled all schools (including Catholic and private) but did not use census data. We suspect the two are comparable and have no reason to believe that Victoria's results would be atypical of Australia as a whole. The loss is of course small and even if young schoolchildren have a slightly worse cognitive environment, older age groups could have a better one.

6.7. Estonia

The last nation listed as showing IQ losses is really a nation in which the data are so confused that their message is unclear. Korgesaar (2013) reports results from three Estonian schools in which school leavers (ages 14–15) took Raven's in 2001, 2005, and 2012. He does not separate male and female trends, but by comparing the full sample and males only, the values for females can be derived.

As Table 4 shows, when this is done, there are several anomalies. Between 2001 and 2005 (5 years), males are supposed to have gained 3.42 IQ points; and then between 2005 and 2012 (7 years) lost 7.68 points. Such a sudden and huge turn around is in itself implausible. Females lost at a steady rate for a total of 8.02 points over the whole 12 years, which in itself is more plausible. However, from 2001 to 2005, the difference between male and female trends is supposed to have been 7.60 points: as if these schools suffered a regime of male dominance (3.42 points gained) and female deprivation (4.18 point lost). Even the 3.66-point gender difference between 2005 and 2012 is extraordinary: males now become completely demoralized (a 7.58-point loss) and females follow suit at a about half the rate (4.02 points lost). Over the 12 years, the graduates from these schools fell from 552

Table 4
Recent Raven's IQ trends in Estonia.

Ages	Test	Years compared	IQ points	Sample and no.	Reference
14–15 males	Raven's	2001–2005	+ 3.42	552 & 411 (both sexes)	Korgesaar, 2013
14–15 females	Raven's	2001–2005	− 4.18		Korgesaar, 2013
F vs. M difference	Raven's	2001–2005	(7.60)		Korgesaar, 2013
14–15 males	Raven's	2005–2012	− 7.68	411 & 304 (both sexes)	Korgesaar, 2013
14–15 females	Raven's	2005–2012	− 4.02		Korgesaar, 2013
F vs. M difference	Raven's	2005–2012	(3.66)		Korgesaar, 2013

Table 5
England and Piagetian tests: decimation at top may foreshadow general IQ decline?
Sources: Shayer et al. (2007); Shayer and Ginsburg (2009). Michael Shayer supplied additional data for values in brackets.

Volume & heaviness (ages 11–12)

Raw score	1975–76	1994	2000–01	2003–04	Percentages
3A & above	(8.30)	Slight increase	(0.85)	(0.53)	Actual
	–	8.73	2.49	1.15	Predicted
2B* & above	28.65	Slight increase	11.65	5.20	Actual
	–	29.56	12.67	8.67	Predicted

Equilibrium in balance (ages 12–13)

Raw score	1976	1994	2006	Percentages
3A and above	20.35	(20.35)	4.8	Actual
	–	–	20.07	Predicted
2B and below	(36.05)	(36.05)	(15.73)	Actual
		–	36.43	Predicted

Pendulum (ages 13–14)

Raw score	1975	1994	2007	Percentages
3A & above	23.80	(23.80)	11.65	Actual
	–	–	27.36	Predicted
2B and below	(29.4)	(29.4)	(18.65)	Actual
	–	–	25.76	Predicted

Volume & Heaviness (ages 11–12): IQ trends. 1975.5–1994: + 0.0293 SD (+ 0.44 IQ); 1975.5–2000.5: − 0.575 SD (− 8.63 IQ); 1975.5–2003.5: − 0.795 SD (− 11.925 IQ); 2000.5–2003.5: − 0.220 SD (− 3.295 IQ).
Equilibrium in Balance (ages 12–13): IQ trends. 1976–1994: + 0.0 SD; 1976–2006: − 0.01 SD (− 0.15 IQ).
Pendulum (ages 13–14): IQ trends. 1975–1994: + 0.0 SD; 1975–2007: + 0.11 SD (+ 1.65 IQ).
Note: 2B = mature concrete level; 2B* = concrete generalization level; 3A = early formal level.

to 304: perhaps something unusual was going on. No one can believe that gender trends among the Estonian population as a whole are represented here.

Korgesaar compares his results to another study. Must, te Nijenhuis, Must, and van Viannen (2009) tested students aged 13–15 in 1n 1997.5 (no. = 361) and 2006 (no. = 670) against the norms of the NIT (National Intelligence Test) that date from 1934.5. The school samples were drawn from the same region and considered comparable. Comparing 1934.5 and 1997.5 the gains for Full Scale IQ amounted to 6.20 IQ points. The data imply 4.73 points because of an arithmetical error but the difference is irrelevant. Note that the gains for 1997.5 to 2006 totaled 2.51 points. This is to say that the rate of gain had tripled in the recent period – gone from 0.98 IQ points per decade (1934.5–1997.5) to 2.95 points per decade (1997.5–2006).

This last period overlaps with the Raven's period of 2001–2012, which showed a loss of 7.90 points per decade (both sexes combined). The NIT (National Intelligence Test) differs from Raven's but is typical of a crystalized IQ test. It includes Vocabulary, Synonyms–Antonyms, Information, Comprehension, Comparisons, Sentence Completion, Analogies, Arithmetic, Computation, and Symbol–Number Correspondence. There is no precedent for fluid and crystalized intelligence to show such different trends at the same time. Estonia should really be classified as a nation in which recent trends are uncertain. Another uncertainty: Estonia has a very high rate of emigration; there is no data that allow us to weight for the possibility of a brain drain.

7. Britain and Norway

Britain has Piagetian data of high quality. For trends on Volume & Heaviness, sixty-nine samples (ages 11–12; school year 7) yielded more than 2000 subjects tested each year (Shayer, Ginsburg, & Coe, 2007). They represented both England and Wales and also state and independent schools. The trends on other tests were based on eight schools with 913 subjects (ages 12–13 & 13–14; school year 8–9). However, their Piagetian scores were compared to national norms on the Cognitive Abilities Test, or the Middle Years Information System Test from the University of Durham (Shayer & Ginsburg, 2009).

7.1. Piagetian trends versus Raven's trends

The Piagetian results below may look irreconcilable with the Raven's trends from Table 2. It showed that British schoolchildren (ages 5–15) gained 6.23 IQ points from 1980 to 2008. However, Flynn (2012a, Table AI8) reveals that when we focus on the older age groups, or the ages that also took the Piagetian tests, the results are quite different. The data for the two tests cover much the same time.

Age 12.37: Raven's + 2.41 IQ points; Piagetian Volume & Heaviness − 11.9 points.
Age 13.33: Raven's + 1.12 IQ points; Piagetian Equilibrium − 0.30 points.
Age 14.25: Raven's - 2.16 IQ points; Piagetian Pendulum + 1.65 points.

Only age 12 poses a problem: the small Raven's gain and the huge Piagetian loss, one that began only after 1993. The 11.9 points lost dictates that the preceding 7.5 years (1985.5 to 1993) would have to show an 11.9 gain to compensate. That is a sudden shift from a rate of gain at 1.6 IQ points per year to a rate of loss at − 1.14 points per year, which is scarcely credible.

There is the possibility that Piagetian tests are measuring something quite different than Raven's, so one could go up and the other down. However, the content of the Piagetian tests include items of ascending

Table 6
Scandinavia: IQ decline without decimation at top.

Denmark (78 test items)				
Raw score	1988	1998	2003–2004	Percentages
52 & above	17.60	20.40	17.40	Actual
	–	20.13	18.00	Predicted

Norway – word similarities (54 items)			
Raw score	1993	2002	Percentages
40 & above	24.90	19.10	Actual
	–	20.03	Predicted

Norway – arithmetic (30 items)			
Raw score	1993	2002	Percentages
22 & above	20.90	14.70	Actual
	–	14.48	Predicted

Norway – figures (36 items)			
Raw score	1993	2002	Percentages
30 & above	25.20	25.90	Actual
	–	26.04	Predicted

Example of calculation (using Denmark).
1988 to 1998 - Actual percentage (1988) for 52 & above is top 24.10.
(1) 0.500–0.241 = 0.259 above median or + 0.703 SDs.
(2) Gains: 1.35 points/15 = + 0.09 SDs.
(3) If that SD gain held over the full curve, those at 52 + would become + 0.703 − 0.09 = + 0.613 SDs above the median, which equals 23.57%.
(4) 0.5000 – 0.2357 = 0.2643 or 26.43% now at 52 & above.
(5) The predicted 26.43% (for 1998) at 52 & above is of course higher than the old actual (1988) of 24.10 thanks to IQ gains. But the true actual for 1998 is only 5.70%.
1988 to 2003–2004 - Actual percentage (1988) for 52 & above is still top 24.10.
(1) 0.500 – 0.241 = 0.259 above median or + 0.703 SDs.
(2) The gains (1988–1998) at 1.35 points are larger than the loss of 1.18 from 1998 to 2003–2004. Thus there is a net gain of 0.1675 IQ points; and 0.1675/15 = + 0.0117 SDs gained.
(3) If that SD gain held over the full curve, those at 52 + would become + 0.703 – 0.0127 = 0.692 SDs above the median, which equals 25.55%.
(4) 0.5000 – 0.2555 = 0.2455 or 24.55% now at 52 & above.
(5) The predicted 24.55% (for 2003–2004) at 52 & above is of course very slightly higher than the old actual (1988) of 24.10 thanks to the net IQ gain. But the true actual for 2003–2004 is only 4.80%.

difficulty that culminate in tasks that require high analytic ability just as Raven's does. For example, in items 13a and 13b of Volume & Heaviness, the teacher tells them the story of Archimedes and the king's crown. They are told that copper is lighter than gold, and that Archimedes first found out the old and the new crown's weights and then their volumes. In 13a they are asked how he got their volume (with a visual hint of him in a full bath), but in 13b they are told that the new crown weighed more than the old, and yet Archimedes still proved there was some lighter metal in it. They have to show what his proof strategy was.

7.2. Piaget and decimation at the top

Setting age 14.25 aside, look at Table 5, which analyzes trends on the three Piagetian tests in detail. The Equilibrium and Pendulum tests are similar. From about 1975–1976 to 2006–2007, both show a huge decimation of the top scorers who could achieve at the early formal level: down from 20.35 to 4.80 and from 23.80 to 11.65% respectively.

The table emphasizes the point by predicting what the top percentages would have been by 2006–2007 if the small overall gains or losses were evenly distributed over the entire curve. The actual values are only 24% of the predicted values for Equilibrium and 43% for Pendulum. On the other hand, if the losses at the top were evenly distributed over the curve, Equilibrium would have lost 0.8357 SDs or 12.54 IQ points at the mean, and Pendulum would have lost 0.4799 SDs or 7.20 IQ points.

The large losses by the top 20 to 24% plus no real change at the mean entail improved performance over the rest of the curve and this was particularly true of those who scored at lower levels, that is, the mature concrete level (2B) or below. In 2006 and 2007, the actual percentages were far less at these levels than the predictions: 15.73 (or 43% of the predicted) for Equilibrium; 18.65 (or 72% of the predicted) for Pendulum. It would be odd if huge gains at the bottom can compensate for huge decimation at the top forever. It seems prudent to anticipate a general decline for Equilibrium and Pendulum.

By 2000–2001, Volume & Heaviness was already in general decline. It would be convenient if it showed decimation at the top some years before (say 1994) as an omen. It does not. However, once its period of decline began in 1995, there were disproportionate losses at the top. Few at age 12 attain the highest Piagetian level (the formal level of 3A and above). The total losses at the mean (− 0.575 and − 0.795 SDs respectively), if projected over the entire curve, would much reduce the top pre-decline percentage of 8.30: the prediction for 2000–2001 is 2.49 and for 2003–2004 is only 1.15%. But the actual percentages were 0.85 and 0.53. These are 34 and 46% of the predicted. In order to see if this is deceptive of the degree of decimation, we include those at the concrete level of generalization and above (2B* and above). At this level, given the overall trend, the top scorers should have gone down to 8.67% and in fact they were 5.20% or only 60% of the predicted. The decimation at the top for Volume and Heaviness is real.

Why does Volume & Heaviness show a general decline over the whole curve when Equilibrium and Balance do not? The former was designed in 1975 mainly to test conservation of matter, volume and weight, usually achieved between 5 and 8 years. Ten (out of 13) questions in V & H assume no processing of information.

Thus they provide evidence of children's development, not at the time of testing, but when they were 5 to 8. The latter two were designed to test the formal thinking that some children develop by the age of 12 to 14 (the time of testing): Pendulum in the context of control of variables and Equilibrium in terms of being able to model compensating variables algebraically.

7.3. Back to conventional tests

Can we find hints in conventional data that mimic the Piagetian trends? During the period of massive IQ gains, a few nations showed higher gains over the bottom half of the curve than over the top half (Flynn, 2012a, pp. 40–43). The only nation outside Scandinavia was Spain. The Spanish sample is both young and local, that is, 7-year olds in Barcelona (Colom, Lluis-Font, & Pueyo, 2005).

Thanks to additional data courtesy of Thomas Teasdale and Jon Sundet, Table 6 analyzes trends in Nordic nations during the most recent period, one in which they actually were in transition (Denmark) or had actually begun to show losses (Norway). There was no new data for Finland and Sweden. The Word Similarities test in Norway shows that the percentage of high scorers fell from 24.90% in 1993 to 19.10 in 2002. The Arithmetic test shows an even large drop, from 20.90% to 14.70. These signal without doubt a decimation of those with high cognitive ability.

However, there is a profound difference between this kind of decimation and that shown by the Piagetian tests. The effects at the top tail of the curve are occasioned by a general decline of performance over the curve as a whole. The cause of the drop is not an atypical decimation of those at the top.

Note the calculations in Table 6. They predict the percentage of high scorers assuming that gains or losses were evenly distributed over the whole curve. The predictions are remarkably accurate. The Word Similarities test in Norway shows a very slight tendency for the predicted percentage of high scorers to be above the actual: 20.03% as compared to 19.10%. This result contrasts with the pattern on the UK Piagetian test of Volume & Heaviness. Once the latter entered a period of losses, within a decade, it showed the actual top scorers at only 46 to 60% of the predicted top scorers. Norway also contrasts with Equilibrium and Pendulum, which although general losses had not begun, showed only 24% and 43% of the predicted top scorers respectively. Denmark, during a period of stable overall performance, joins Norway with no atypical decimation at the top.

There was one period some 40 to 50 years ago in which Norway showed a shift suggestive of the Piagetian pattern. Rist (1982) analyzed the cohorts from 1958 and 1967. It was a decade of no real change at the median hidden in the long period of gains that extended from 1954 to 1993. His Table 6 uses raw scores to separate three groups for General Level of Intelligence (GLI - it merges the results from the three tests). The lowest level of raw scores showed little change (8.9 up to 10.1%) but the top scorers fell off from 45.9 to 35.5% or only 77% of what we would expect. Unfortunately, for the thesis that this should trigger general losses, overall performance improved slightly over the next decade (1968–1977).

7.4. How do Piagetian tests differ?

Despite the large decrease in top scorers in Norway, its data do not really duplicate the enormous decimation of top scores that two of the Piagetian tests exemplify. There is no sign that Norway's top scorers were especially susceptible to factors that reduce scores at high levels. At present, we know of no conventional test that shows results of this sort. Why should this be so?

The obvious explanation is that Piagetian tests include numerous items that are cognitively demanding to a very high degree. They are unique in that they distinguish actual levels of competence. You can either get the tough items right that test for competence at the formal level or you cannot. When raw scores are used, conventional tests can detect an atypical drop at the top. However, their items are chosen to discriminate equally over the entire curve and having too many tough ones at the top would be "wasted". At the top, Piagetian tests ask you to perform a feat whose difficulty is fixed: you can either do a triple somersault (without a trampoline) or you cannot. Scoring in the top 2.27% (an "IQ" of 130) of gymnasts will not do. In sum: at one time the best of Britons (aged 12–14) could cope with items on the formal level and blended into a smooth curve of performance. Now these items are beyond many of them and register as a huge decimation of high scorers.

Piagetian gains at the bottom of the curve should not be dismissed as simply a phenomenon that offsets losses at the top. Consider the British results for Equilibrium and Pendulum. The decimation of top scorers means that by the age of 12 to 14, fewer British schoolchildren attain the level of formal operations. This means that fewer could think in terms of abstractions (without concrete examples), which limits their capacity for deductive logic and systematic planning. However, the fact that these losses are made up by gains over the rest of the "curve" means that far more of them are at the concrete generalization level. They are better at on the spot thinking (e.g. in playing demanding computer games). Their understanding of the physical world is limited to simple causation between two variables, but they can draw inferences from observations to make generalizations.

We have seen that the results for Piaget Volume & heaviness have another disturbing implication. Although the subjects were 12.27 years old when assessed, the profound losses they show at the concrete generalization level and below could hardly have suddenly emerged solely at that age. These children must have been well behind earlier cohorts from at least the age of 5 to order to register such losses. However, we need not appeal to the peculiarities of Piagetian tests to make the point. Table 7 uses UK Raven's data. It gives good real-world evidence of how quickly losses can fade way as you go down the age scale. It shows both why Nordic losses at age 18 do not imply any losses before age 16; and why Volume and Heaviness losses at age 12 imply losses into early childhood.

8. Discussion

Nations that are candidates for IQ decline give evidence that ranges from beyond dispute to fragmentary. The strongest comes from Scandinavia as a whole, Britain (Volume and Heaviness), and Germany (spatial). Nordic data cover vocabulary, similarities, analogies, shapes, metal folding, number series, geometrical figures, and letter matrices. The pivotal year for all of these nations seems to be about 1995. Data from France and Australia await further studies but cannot be dismissed. The US is strange. All ages plow on unaffected as yet. Black gains are larger than white but white gains are still robust (Flynn, 2012a). The quick developing world (Korea) and the slow developing world will of course continue to gain for some time.

These trends come as no surprise. Flynn has argued that industrialization may eventually pay diminishing returns in developed

Table 7
Using Raven's data to assess how much IQ trends over time can alter (between gains and losses) with age.
Sources: Table 1, Table 5, and Flynn (2012a), p. 232.

UK Raven's: 1980–2008 (28 years). Selected to assess possible Nordic trends with age			
Actual age	"Nordic age"	IQ trend	IQ difference with oldest age
14.25	18	− 2.26 (loss)	–
13.33	17	+ 1.12 (gain)	3.38
12.37	16	+ 2.41 (gain)	4.67
11.21	15	+ 3.78 (gain)	**6.04**
10.25	14	+ 6.22 (gain)	8.48
Finland: 1997–2009 (12 years but trend projected over 28 years)			
Actual age	–	IQ trend	–
18	–	− **6.99** (loss)	–
Denmark: 1998–2003.5 (5.5 years but trend projected over 28 years)			
Actual age	–	IQ trend	–
18	–	− **6.05** (loss)	–
Norway: 1993–2002 (9 years but trend projected over 28 years)			
Actual age	–	IQ trend	–
18	–	− **6.07** (loss)	–
Conclusion: In all cases, losses may have become absent by age 15			

UK Raven's: 1980–2008 (28 years). Selected to assess possible UK volume & heaviness trends with age			
Actual age	"V & H age"	IQ trend	IQ difference with oldest age
14.25	12	− 2.26 (loss)	–
8.25	6	+ 13.66 (gain)	**15.92**
UK volume & heaviness: 2000.5–2003.5 (3 years but trend projected over 28 years)			
Actual age	–	IQ trend	–
12	–	**32.95**	–
Conclusion: even at age 6, losses would be substantial and significant at lower ages			

nations. Until very recently, we have enjoyed a more favorable ratio of adults to children in the home, more and better schooling, more cognitively demanding jobs, and better health and conditions of the aged. These caused large IQ gains for several generations. However, the same factors can turn from positive to mixed or even negative. The number of children in the home has reached a minimum (and indeed there are more solo-parent homes), middle class parents have used up the tricks that make the pre-school environment cognitively enriching, we appear to have reached a limit in terms of enhanced schooling and the number we keep in school into adulthood, the economy may be producing fewer cognitively demanding jobs in favor of more service work; however, we may continue to improve the health of the aged.

Thanks to the Scottish data (Staff et al., 2014) and Table 7, we know that a decline at one age does not entail a decline at another. The worlds that encompass our lives, early home environment, schools and teenage subculture, university, the world of work, and the conditions of the aged evolve to some degree independently of one another. Even during the period of gains, they never marched in lock step (never gave the same gains from age to age), but at least they were all positive. During a period of "decline" there may be a mix. We would expect ambiguous results and a pattern like the Netherlands: stability among preschoolers, perhaps some losses at school, eventual losses among adults (here The Netherlands may be lagging a bit), and gains among the aged.

Against this, it is always possible that the Scandinavian data represent a pattern that will spread to all ages and become typical of much of the developed world. But remember that Scandinavia gives results only for 18-year-olds and it might show gains for all ages other than the late teens (unlikely for adults, likely for the aged).

Still, there is no reason for optimism. If the economists are correct, adult gains are due to reverse throughout the developed world. The UK Volume & Heaviness test suggests losses that cover at least ages 5 to 12. Shayer speculates about the evolution of British society over the relevant period (since 1989 – when the subjects were 5 or 6). Children drifted away from formal toward concrete thinking. They became more and more immersed in modern visual and aural electronic culture. More time (four to five hours a day, more on weekends) spent on TV, computer games, and cell-phones, all of which decrease their attention span. A hypothesis: When Volume and Heaviness was first used in the CSMS National survey in 1975–6 there was a boy/girl IQ difference of 0.50 SD. At the time, this was attributed to differential play patterns between the ages of 4 to 8. These were unaltered until 1994. After that, thanks to the new electronic culture, the way in which boys and girls spent their free time gradually became the same. As a result, the IQ gender gap steadily decreased. By 2002 it disappeared and in 2003, boys and girls began to suffer IQ decline in tandem (Shayer et al., 2007, Fig. 2).

The Piagetian results are particularly ominous. Looming over all is their message that the pool of those who reach the top level of cognitive performance is being decimated: fewer and fewer people attain the formal level at which they can think in terms of abstractions and develop their capacity for deductive logic and systematic planning. They also reveal that something is actually targeting that level with special effect, rather than simply reducing its numbers in accord with losses over the curve as a whole. We have given our reason as to why the Piagetian tests are sensitive to this phenomenon in a way that conventional tests are not.

Massive IQ gains over time were never written in the sky as something eternal like the law of gravity. They are subject to every twist and turn of social evolution. If there is a decline, should we be too upset? During the 20th century, society escalated its skill demands and IQ rose. During the 21st century, if society reduces its skill demands, IQ will fall. Nonetheless, no one would welcome decay in the body politic, or among the elite who at present represent our best thinkers. Although it might be argued that the character of the electorate will be enhanced if it contained fewer lawyers and more plumbers and service workers.

It is always possible that our schools and universities will graduate more young people who read and become more critically astute. This in itself would put a limit on IQ losses on Vocabulary, Information, and most Verbal tests, and on accepting the stereotypes that cloud moral reasoning and political prudence. Capitalizing on a people's intelligence, rather than worrying about their intelligence, is the most important thing (Flynn, 2010, 2012b, 2012c, 2016b). Ignorance cripples every generation. This is not to deny that if dysgenic mating alters the potential of the human brain over a century, we should be quiescent. But the solution to that problem would take us far afield (Flynn, 2013).

References

Bakhiet, S. F. A., Baraket, S. M. R., & Lynn, R. (2014). A Flynn effect among deaf boys in Saudi Arabia. *Intelligence, 44*, 75–77.

Bandeira, D. R., Costa, A., & Arteche, A. (2012). The Flynn effect in Brazil: Examining generational changes in the Draw-a-Person and in the Raven's coloured Progressive Matrices. *Revista Latinoamericana de Psicologia, 44*, 9–18.

Colom, R., Flores-Mendoza, C. E., & Abad, F. J. (2007). Generational changes on the Draw-a-Man tests: A comparison of Brazilian urban and rural children tested in 1930, 2002, and 2004. *Journal of Biosocial Science, 39*, 79–89.

Colom, R., Lluis-Font, J. M., & Pueyo, A. A. (2005). The generational intelligence gains are caused by decreasing variance in the lower half of the distribution: Supporting evidence for the nutrition hypothesis. *Intelligence, 33*, 83–91.

Cotton, S., Kiely, P., Crewther, D., Thompson, B., Laycock, R., & Crewther, S. (2005). A normative and reliability study of the Raven's Coloured Matrices for primary school aged children from Victoria, Australia. *Personality and Individual Differences, 39*, 647–659.

Dutton, E., & Lynn, R. (2013). A negative Flynn Effect in Finland, 1997–2009. *Intelligence, 41*, 817–820.

Dutton, E., & Lynn, R. (2015). A negative Flynn Effect in France, 1999 to 2008–9. *Intelligence, 51*, 67–70.

Florida, R. (2014). *Where the good and bad jobs will be, 10 years from now. Vol. 25*, City Lab2014.

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

Flynn, J. R. (2009). *What is intelligence? Beyond the Flynn Effect.* Cambridge University Press.

Flynn, J. R. (2010). *The torchlight list: Around the world in 200 books.* Wellington, New Zealand: AWA Press.

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

Flynn, J. R. (2012b). *Fate and philosophy: A journey through life's great questions.* Wellington, New Zealand: AWA Press.

Flynn, J. R. (2012c). *How to improve your mind: Twenty keys to unlock the modern world.* London: Wiley-Blackwell.

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

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

Flynn, J. R. (2016b). *The new torchlight list: In search of the best modern authors.* Wellington, New Zealand: AWA Press.

Harrison, T. L., Shipstead, Z., & Engle, R. W. (2014). *Why is working memory capacity related to matrix reasoning tasks?* Psychonomic Society, Inc.2014.

Kong, A., Frigge, M. L., Thorleifsson, G., Stefansson, H., Young, A. I., Zink, F., ... Stefansson, K. (2017). Selection against variants in the genome associated with educational attainment. *Proceedings of the National Academy of Sciences, 114*, E727–E732.

Korgesaar, M. (2013). *Flynni Efekti esinemine ee sti abiturientide seas Raveni Testi põhjal [presence of Flynn effect among Estonian school-leavers on raven matrices].* Seminaritöö: Tartu Ülikool, Psühholoogia instituut.

Must, O., te Nijenhuis, J., Must, A., & van Viannen, A. E. M. (2009). Comparability of IQ scores over time. *Inteligence, 37*, 25–33.

Petermann, F., & Petermann, U. (Eds.). (2008). *Hamburg-Wechsler-Intelligenztest für Kinder IV (HAWIK-IV)*. Bern: Verlag Hans Huber.

Pietschnig, J., & Gittler, G. (2015). A reversal of the Flynn effect for spatial perception in German-speaking countries: Evidence from a cross-temporal IRT- based meta-analysis. *PLoS One, 5*, e14406.

Pietschnig, J., Voracek, M., & Formann, A. K. (2010). Pervasiveness of the IQ rise: A cross-temporal meta-analysis. *Intelligence, 53*, 145–153.

Rist, T. (1982). *Det intellektuelle prestajonsnivaet I befolkningen sett I lys av den samfunns-messige utviklinga [The level of the intellectual performance of the population seen in the light of developments in the community].* Oslo: Norwegian Armed Forces Psychology Service.

Rönnlund, M., Carlstedt, B., Blomstedt, Y., Lars-Göran Nilsson, L.-G., & Lars Weinehall, R. (2013). Secular trends in cognitive test performance: Swedish conscript data 1970–1993. *Intelligence, 41*, 19–24.

Shayer, M., & Ginsburg, D. (2009). Thirty years on – A large anti-Flynn effect? (II): 13- and 14-year-olds. Piagetian tests of formal operations norms 1976–2006/7. *British Journal of Educational Psychology, 79*, 409–418.

Shayer, M., Ginsburg, D., & Coe, R. (2007). Thirty years on – A large anti-Flynn effect? The Piagetian test volume & heaviness norms 1975–2003. *British Journal of Educational Psychology, 77*, 25–41.

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

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*, 349–362.

te Nijenhuis, J. (2013). The Flynn effect, group differences and g loadings. *Personality and Individual Differences, 55*, 224–228.

te Nijenhuis, J., Cho, S. H., Murphy, R., & Lee, K. H. (2012). The Flynn effect in Korea: Large gains. *Personality and Individual Differences, 53*, 147–151.

Teasdale, T. W., & Owen, D. R. (2008). Secular declines in cognitive test scores: A reversal of the Flynn Effect. *Intelligence, 36*, 121–126.

Weiss, L. G., Gregoire, J., & Zhu, J. (2016). Flaws in Flynn Effect research with the Wechsler Scales. *Journal of Psychoeducational Assessment, 34*, 411–420.

Wicherts, J. W., Dolan, C. V., Oosterveld, P., van Baal, G. C. V., Boomsma, D. I., & Span, M. M. (2004). Are intelligence tests measurement invariant over time? Investigating the nature of the Flynn effect. *Intelligence, 32*, 509–537.

Woodley of Menie, M. A., & Fernandes, H. B. F. (2015). Do opposing secular trends on backwards and forwards digit span evidence the co-occurrence model? A comment on Gignac (2015). *Intelligence, 50*, 125–130.

Woodley of Menie, M. A., & Meisenberg, G. (2013). In the Netherlands the anti-Flynn effect is a Jensen effect. *Personality and Individual Differences, 54*, 872–876.

Woodley of Menie, M. A., Peñaherrera, M., Fernandes, H. B. F., & Figueredo, A. J. (2017). What causes the anti-Flynn Effect? A data synthesis and analysis of predictors. *Evolutionary Behavioral Sciences*. http://dx.doi.org/10.1037/ebs0000106 (in press).