Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Intelligence 31 (2003) 461–471
INTELLIGENCE

The secular rise in IQs: In Estonia, the Flynn effect is not a Jensen effect

Olev Must*, Aasa Must, Vilve Raudik

Department of Psychology, University of Tartu, Tiigi 78, Tartu 50410, Estonia

Received 16 January 2001; received in revised form 11 November 2002; accepted 14 November 2002

Abstract

This study investigated the secular rise in IQ scores over a 60-year period in 12- to 14-year-old Estonian schoolchildren. In 1934/1936, Juhan Tork adapted the U.S. National Intelligence Test for Estonia and administered it to more than 6000 schoolchildren. We administered the same test to 449 students in 1997/1998 and compared the results of 381 of these with a carefully matched sample of 307 from the testing in the 1930s. We found a rise of nearly 1 S.D. on subtests using basic information-processing algorithms such as Comparison and Symbol–Number, but only 0.50 S.D.s on verbal subtests such as Sentence Completion and Concept Comprehension. The secular gains were most pronounced on the low *g*-loaded subtests. In two compared age groups of children, the rank order correlations between the secular changes on the various subtests and the rank of those subtests on the *g* factor are negative and nonsignificant, the mean $r_s = -.40$ (one-tailed $P = .13$). As such, these results supported Rushton's [Pers. Individ. Differ. 26 (1999) 381] finding that the secular rise over time is not occurring on the *g* factor. In Estonia, the Flynn effect is not a Jensen effect.

Keywords: Flynn effect; Jensen effect; *g* Factor

1. Introduction

The question has arisen whether the secular increase in IQ scores over time is occurring on the *g* factor, that is, whether the "Flynn effect" is a "Jensen effect." The Flynn effect refers to

* Corresponding author. Tel.: +372-7-375-912; fax: +372-7-352-900.
E-mail address: olev@psych.ut.ee (O. Must).

0160-2896/03/$ – see front matter
doi:10.1016/S0160-2896(03)00013-8

the repeated demonstration that the populations of several countries have increased in average IQ by about 3 points a decade over the last 60 years (see Flynn, 1984, 1987, 1999; it is also sometimes known as the Lynn–Flynn effect; see Lynn, 1982). Similarly, the Jensen effect refers to the repeated finding that the vector of a test's *g* loadings is the best predictor of that test's correlation with a variety of variables, including not only scholastic and workplace performance, but also brain size, brain pH, brain glucose metabolic rate, average evoked potential, reaction time, and other physiological factors (see Jensen, 1980, 1998). The question is whether these two effects are related.

Rushton (1999) was the first to find evidence that the secular rise in IQ did not appear on *g*, and he concluded that the Flynn effect was not a Jensen effect. Rushton carried out a principal components analysis of the secular gains in IQ on the Wechsler Intelligence Scale for Children (WISC-R and WISC-III) from the United States, Germany, Austria, and Scotland, along with *g* loadings from the standardization samples, inbreeding depression scores from cousin marriages in Japan, and Black/White IQ difference scores from the United States. The results were the following: (1) the IQ gains on the WISC-R and WISC-III formed a cluster, showing that the secular trend is a reliable phenomenon, but (2) this cluster was *independent* of the cluster formed by *g*-factor loadings, inbreeding depression scores, and Black/White differences. Rushton's analysis showed that the secular increase in IQ was unrelated to *g* and other heritable measures.

Rushton's (1999) results, however, were contradicted by a subsequent study carried out in Spain. Colom, Juan-Espinosa, and Garcia (2001) reported a positive correlation (r=.78; P < .05) between *g* and the amount of generational change in two successive standardizations of the Spanish Differential Aptitude Test across 16 years. There were 10 samples of males and females for each of five subtests (Verbal Reasoning, Space Relations, Numerical Ability, Mechanical Reasoning, and Abstract Reasoning). However, there were ambiguities in the study of Colom et al. that raised questions about its generality. For example, 5 of the 10 samples showed a generational *decrement* (their Table 1).

The controversy is even more deep rooted because from the first discovery of the effect there have been contradictory findings with scores on some tests indicating a rise in scores, others indicating a fall, and still others indicating no change (Neisser, 1998). For example, Raven (2000) concluded that whereas convincing gains were apparent in "eductive ability" (abstract reasoning, as measured on the Raven's Progressive Matrices), the gain was ambiguous in "reproductive ability" (the ability to reproduce previously learned information, as measured on the Mill Hill Vocabulary Scale). Still, other data indicate that in some modern societies, such as Sweden, the IQ gain has been minimal or even absent (Svensson, Emanuelsson, & Reuterberg, 1997).

Questions concerning the reliability of the secular increase in test scores, and whether it occurs on the *g* factor, can only be answered with new research. Accordingly, the present paper brings new data from Estonia to bear on the issues. These Estonian data (from 1934 to 1936) are some of the oldest available for the comparison of two samples. They were gathered by Juhan Tork, an Estonian educator who was supported by the Estonian Ministry of Education. As part of his doctoral dissertation, Tork (1940) adapted the National Intelligence Test (Terman, 1921) from English to Estonian, and administered it to 6000 Estonian schoolchildren. The original tests, plus much background information, can be accessed in the Historical Archive of Estonia. Although the language and the content of some items were out of date, we adapted the test with minimal changes, administered it to new samples, and so gained appropriate data for comparative purposes.

2. Methods

The National Intelligence Test consists of 10 subtests: Arithmetic (16 items requiring a solution for one unknown quantity); Computation (22 items requiring addition, subtraction, multiplication, and division of both integers and fractions); Sentence Completion (20 items requiring filling in missing words to make sentences understandable and correct); Information (40 items requiring picking the correct answer from an array of questions such as "What is the first month of the year?"); Concepts (24 items requiring selecting two characteristic features from among those given such as "Cat: sphere, claws, eating, eyes, mouse?"); Vocabulary (40 items requiring knowledge about the qualities of different objects such as "Are books used in school?"); Synonyms–Antonyms (40 items requiring evaluation of whether the words presented mean the same or opposite); Analogies (32 items requiring transferring the relation between two given words to other presented words); Symbol–Number (120 items in which the correct digit must be assigned to a presented symbol from a key); and Comparisons (50 items requiring same or different judgments about sets of numbers, family names, and graphic symbols presented in two columns).

When Tork (1940) adapted the National Intelligence Test, he made changes to every subtest but one: Symbol–Number. Minimal non-content-related changes were made in the Arithmetic, Computation, and Comparison subtests. The other subtests got new Estonian specific content. In the present replication study, additional adaptations were minimal. Some words were changed and some facts in the Information subtest were brought up to date (e.g., to include World War II).

In comparative studies such as this one, it is important to match groups carefully. Tork's (1940) sampling was not fully representative of the Estonian population, being regionally biased, with 75% of the test takers coming from one county, Tartumaa, and from one city, Tartu. Moreover, Tartu was not a typical Estonian city for it contained Estonia's only university. (One of Tork's main findings was that test scores varied from school to school.) The age/grade sampling was mainly of 12- to 14-year-olds in Grades 5 and 6, the final grades of primary school. He also found that each grade passed in the school added 1 S.D. to the test score (with age held constant). Our replication had to consider these facts and that the age/grade relationship had changed by two school years. Students in the 1990s had attended school for 2 years longer than had their same-age counterparts in the 1930s. Hence, it would be incorrect to ignore the differences in educational impact and compare students only by age. Taking school grade as the base for comparison would also be insufficient for there have been substantial changes to the educational system in 60 years, including the content of education, the school entry age, and the criteria for promoting students from one grade to the next. The age structure per grade was more heterogeneous in the 1930s than it was in the 1990s. Thus, we decided to compare only *modal student groups*. The modal pattern of the age/grade relationship in the 1930s was as follows: 12- to 13-year-olds in the 5th grade, and 13- to 14-year-olds in the 6th grade. In 1997/1998, these age groups were attending the 7th and 8th grades, respectively.

2.1. The sampling strategy

Tork's (1940) main aim was to describe the students graduating from primary school in the 5th and 6th grade. By the 1990s, the same age group was in the 7th and 8th grade. The pilot research revealed that the mean age of students attending these grades was 13.2 years. This was the leading criterion for finding test booklets that had been completed in 1934/1936. Additional criteria for matching samples that we applied in the archive search were equal gender and regional balance, rural students (Tartu county), students from Tartu city, and from Tallinn (Estonia's capital, the most urban area). Comparative data were gathered by grades from the same regions and in equal proportions. Testing was carried out in 1997/1998 following the same procedure as the original work. This initial sampling was then corrected by the modal student model. This procedure diminished age diversity and matched the samples better so as to examine the Flynn effect/Jensen effect. In order to examine whether the secular rise occurred in different age groups, the modal sample was also divided into a younger and an older group.

3. Results

The key issue is whether students in the 1990s achieved higher scores than those in the 1930s. Keeping in mind the above-mentioned concept of a modal student, and analysing only

Table 1
Changes in the IQ test scores: the Flynn effect, 1934/1936–1997/1998

	1934/1936 younger		1997/1998 younger		1934/1936 older		1997/1998 older		Effect size *d*	
	N = 72		*N* = 240		*N* = 235		*N* = 141		Younger	Older
	M	S.D.	*M*	S.D.	*M*	S.D.	*M*	S.D.		
Subtest										
Arithmetic	13.89	4.07	14.16	4.50	17.85	4.17	15.15	4.42	0.06	− 0.63
Computation	21.89	4.24	22.02	5.64	25.23	5.30	22.67	4.81	0.02	− 0.51
Sentence Completion	23.11	5.94	29.12	6.22	28.46	6.07	31.05	5.45	0.99	0.45
Information	19.22	5.29	18.72	5.69	27.11	5.96	20.02	4.90	− 0.09	− 1.30
Concepts	33.53	7.95	38.55	7.44	36.23	8.26	40.15	6.06	0.65	0.54
Vocabulary	23.08	4.98	24.22	6.72	26.44	5.42	25.89	5.97	0.19	− 0.10
Synonyms–Antonyms	19.93	13.08	29.34	8.09	28.67	12.15	29.82	8.38	0.86	0.11
Analogies	10.75	4.65	16.70	6.25	14.81	5.76	18.27	5.37	1.08	0.62
Symbol–Number	21.50	5.02	29.60	5.97	25.13	6.83	30.78	4.88	1.46	0.95
Comparisons	23.17	6.40	32.50	8.66	28.30	8.59	33.86	8.53	1.22	0.65
Test in all	210.07	39.16	254.94	43.54	258.24	40.67	267.66	37.21	1.08	0.24

test performance, the results are shown in Table 1 and Fig. 1. It is apparent that overall performance rose on the National Intelligence Test over the 60 years, and that the changes varied by subtest. While most subtests showed increases, some showed decreases. For example, the subtests that needed computation skills and mathematical thinking were not performed better at the end of 1990s than they were 60 years earlier, and scores on the Information subtest had declined (although the items in that subtest are about the Estonian history, cultural figures, and living conditions of the 1930s, which present students are less aware of). Moderate gains were apparent in verbal subtests and the most impressive gains of all were in the Symbol–Number and Comparison subtests.

The test score gains were bigger in the younger than in the older group. Age, school grade, gender, and regional location of the school are classic background variables in IQ research, but still there are few data demonstrating variation of the Flynn effect in those dimensions. Our data collection and sampling procedures allowed us to perform an ANOVA to check influences of independent background variables and their interactions on the test results, and so we examined the effects of age, gender, testing period, and region. Table 2 describes the main effects as well as the two-way interactions. Test performance was substantially influenced by the testing period, age group, and location of the school. There were also two significant interactions: the interaction between gender and testing period, and between age and testing period.

The period of testing (1930s/1990s) had a substantial effect on test performance. The secular gain in IQ was found in both males and females, in both age groups, and in all three

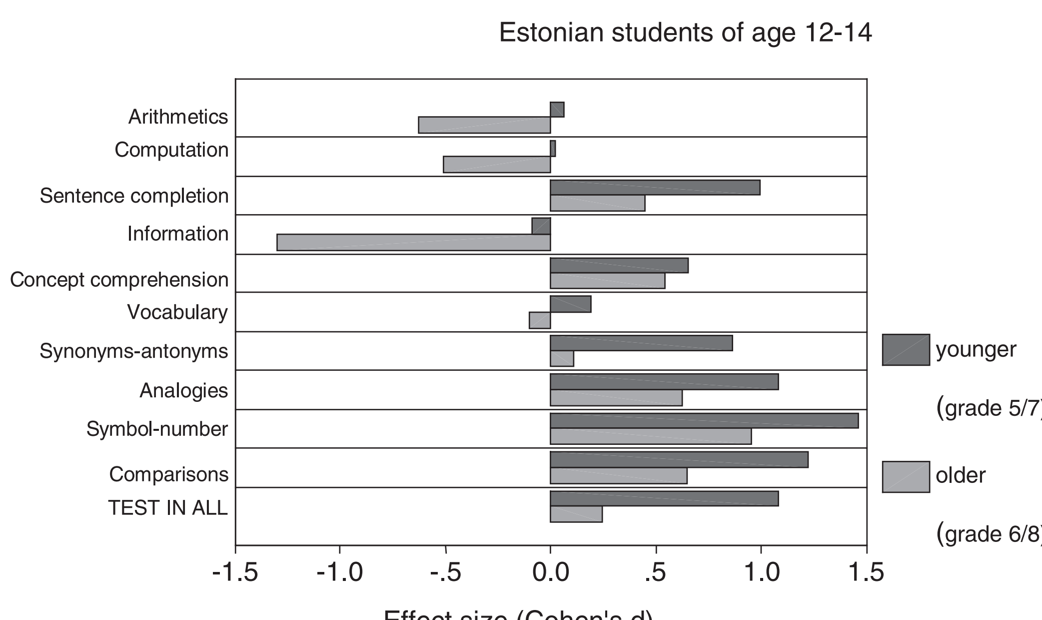

Fig. 1. Changes in the IQ test scores: the Flynn effect, 1934/1936–1997/1998.

Table 2
Dependence of test performance on time of testing,[a] location of school,[b] gender,[c] and age group[d] (ANOVA, main effects, and two-way interaction effects)

Source of effect	Degrees of freedom (*df*)	*F*	Significance of effects (*P*)
Model	15	1843.489	.000
Region	2	7.617	.001
Age group	1	58.543	.000
Time	1	54.888	.000
Gender	1	1.705	.192
Time × Gender	1	4.152	.042
Time × Age	1	18.536	.000
Time × Region	2	0.364	.695
Region × Gender	2	1.543	.215
Region × Age	2	1.663	.190
Gender × Age	1	0.852	.356

[a] Time of test performance: two groups—original (1934/36) and replication (1997/98).
[b] Location of school: three regional groups—Tallinn, Tartu city, county of Tartu.
[c] Gender: two groups—girls and boys.
[d] Age group: two groups of modal students—younger (Grade 5/7, 12–13 years old), older (Grade 6/8, 13–14 years old).

regions. The age effect was also significant, with the test performance of the older modal group better than that of the younger group, irrespective of the period of testing, school location, or gender. The school location also had an effect on the test performance. In the 1930s, as well as in the 1990s, the students in Tartu city had the highest scores, followed by students in Tallinn, and then the students from the countryside. There was, however, no

Table 3
Principal component of test and factors after varimax rotation

Subtest	Original	1934/1936		Replication	1997/1998	
	Principal component	F1	F2	Principal component	F1	F2
Information	.885	.786	.427	.810	.798	.253
Sentence Completion	.845	.821	.313	.741	.809	.105
Analogies	.809	.636	.501	.782	.700	.357
Arithmetic	.782	.603	.500	.639	.706	.079
Vocabulary	.780	.784	.254	.743	.710	.266
Concepts	.735	.793	.166	.597	.496	.332
Symbol–Number	.669	.315	.697	.509	.094	.807
Synonyms–Antonyms	.669	.610	.301	.647	.567	.314
Comparisons	.653	.324	.657	.613	.237	.774
Computation	.623	.172	.812	.689	.462	.560
Eigenvalue	5.62	3.91	2.54	4.66	3.63	2.05

Table 4
Correlations of subtests (Pearson's correlation coefficient)

	Arithmetics	Computation	Sentence Comprehension	Information	Concept Comprehension	Vocabulary	Synonyms–Antonyms	Analogies	Symbol–Number	Comparisons
Arithmetic		.489	.648	.682	.473	.529	.504	.573	.479	.427
Computation	.412		.434	.478	.315	.400	.336	.478	.438	.397
Sentence Comprehension	.492	.355		.756	.654	.655	.514	.642	.480	.429
Information	.484	.476	.601		.608	.732	.551	.667	.518	.533
Concepts	.230	.267	.440	.468		.563	.432	.570	.451	.378
Vocabulary	.404	.458	.534	.627	.347		.458	.580	.403	.430
Synonyms–Antonyms	.379	.353	.469	.405	.339	.394		.476	.333	.438
Analogies	.451	.533	.503	.621	.423	.518	.454		.525	.490
Symbol–Number	.230	.341	.254	.264	.313	.272	.306	.309		.444
Comparisons	.242	.479	.296	.420	.296	.388	.330	.401	.411	

Above the diagonal are data from 1934/1936.
Below the diagonal are data from 1997/1998.

evidence of a linear relationship between the urbanisation level of the school region and test performance in 1990s.

Several interactions among the variables were significant. These included those of gender and time, and age and time. Gender differences in test performance have changed in the course of time. In the 1930s, the test performance of the boys was minimally better than that of the girls, whereas in the 1990s, the test performance of the girls was better. In the younger group, the differences between the test performance are larger after 60 years than in the older modal group. This trend is the same for boys and girls, both urban and rural students.

Tork (1940) was mainly interested in the total test score and did not carry out any analyses at the subtest level. The distinction between fluid and crystallized intelligence, an important topic in modern discussions, did not exist in his time. However, because all the data were available, we were able to key in the subtest scores (which correlated substantially with each other) and carry out a factor analysis of his data as well as of our own.

The factor analysis of Tork's (1940) data showed the presence of one strong factor (accounting for 56% of the variance) (Table 3). A factor analysis of our replication data shows this first factor is slightly weaker (accounting for only 47% of the variance). The correlation between subtests had also changed somewhat, decreasing some of the former integrity and mutual relations of subtests. Of 45 possible subtest intercorrelations, only 4 were higher in our replication study, whereas 41 were smaller (Table 4). The 1930s data revealed only one factor with an eigenvalue of over 1, whereas the 1990s data revealed two factors with eigenvalues over 1. In the replication study, the first factor can be described as a Verbal Comprehension factor and the second factor is described by the high factor loadings of the Symbol–Number, Comparison, and Computation subtests.

A comparison of the first principal components does not support the hypothesis that the secular changes in IQ are on the *g* factor. The congruence coefficient of the first principal component across the 60 years is +.996. This means that the general factors of the two matrices are virtually identical. Thus, the differences found in the correlational matrices do not mean there is a change in the structure of *g*. This conclusion is supported by the fact that in both age groups of children, the rank order correlations between the secular changes on the various subtests and the rank of those subtests in the factor structure are negative and nonsignificant. The mean Spearman's rank order correlation coefficient is $-.40$ (one-tailed $P=.13$) The Flynn effect is mostly due to the least *g*-loaded tests in this battery, as they were in Rushton's (1999) earlier study.

4. Discussion

The research presented here adds new information to the ongoing debate about the secular rise in IQ scores, also known as the Flynn effect (Neisser, 1998). First, these Estonian data corroborate previous studies showing that IQ test scores increase over time, and does so using the National Intelligence Test, which has not previously been used in such research and yet which is older than either the Raven's Progressive Matrices or the Wechsler scales and with a different theoretical background and set of measurement properties. Importantly, these

Estonian data also demonstrated that the rising IQ test scores has occurred in a less developed country. Some critics have argued that the literature is limited mainly to data from economically developed countries (Rodgers, 1999). Thus, our data support the universal nature of the effect. In the 1930s, Estonia was a typical agricultural country, established as independent only 10–15 years previously. School education and the basic pragmatic intellectual skills (computation) were highly valued. It is possible to consider the Estonia of those years as an example of a traditional society (Inglehart, 1995). Historical events beginning in the 1930s—most obviously including World War II, 50 years of Soviet occupation, the reestablishment of independence, and now the global information age—radically changed that society, its values, and emphases. It could be imagined that the Estonian people would invest their cognitive energy into short-term memory and quick algorithm-based information processing rather than in abstract problem solving. To answer Rodger's (1999) question of whether the Flynn effect is an age, cohort, or period effect, it seems that in Estonia it is a period effect.

The results reported here both confirm and differ from earlier findings. Similar to earlier findings (Flynn, 1999), Estonian children perform worse in mathematical subtests. Although we found a gain in verbal subtests (Analogies, Sentence Completion), the most impressive gain was in Symbol–Number and Comparisons subtests. The task demands in those subtests are for short-term memory and attention, as well as effectiveness in algorithm-related information processing (Cattell, 1987; Horn & Noll, 1997). The technical information-processing aspect seems to be especially important: the items of the subtests could easily be solved with some aid (calculator, word-processing programmes). If we were to agree with Flynn (1999) about the influence of society's cognitive demands on IQ test results, then this interpretation may be appropriate.

One difference from previous studies of the Flynn effect is our finding about gender differences. Like the Swedish data from Svensson et al. (1997), our study revealed that the Flynn effect is stronger in girls than in boys. A social interpretation seems to be appropriate here. In the 1930s, women's rights were not a topic for discussion in Estonia. Traditional societies are rather masculine. But the Soviet system changed the situation. Socialist regimes typically emphasized women's rights and schooling, and in socialist societies, education has been relatively favorable to women, compared to other spheres of social life (Jacobs, 1996). In Estonia, nowadays, girls manage better than before with the tasks used in tests and in educational evaluations.

Our finding that the Flynn effect appeared more strongly in the younger modal group than in the older group needs additional attention. The mean test score of the present 7th grade students was 1.14 S.D.s higher than that of their counterparts in the 1930s but the older group scored only 0.23 S.D.s higher. Several explanations are possible, including earlier maturity. The beginning of intensive socialisation has moved to an earlier age and the plurality of socialisation agents cause earlier development. This might be the reason why the influence of one school year was bigger in the 1930s than it was in the 1990s. The differences between grades in the 1930s were in general very large. That is why Tork (1940) focused on age–grade covariation. The test performance of the 6th grade students of Tork's time was 1.2 S.D.s better than the 5th grade students, whereas the present shift from 7th to 8th grade was only 0.29 S.D.s.

Some qualitative changes may also be apparent. If we proceed from the test used, we can conclude that the intellectual functioning of Estonian students in the period under discussion is acquiring a new structure because the subtests evaluating the different aspects of intellectual functioning correlate differently. In the 1930s, 10 subtests formed one clear factor, whereas in the 1990s, the explanatory power of the first principal component was slightly smaller and a second factor could be distinguished. Thus, there is ground for speculating about the emergence of relatively independent skills or the ability to process information quickly and effectively rather than overly abstractly. In this context, it is worth noting that Detterman and Daniel (1989) found the *g* factor accounts for less variance in more intelligent groups of subjects. Interpolating this finding to the context of rising test scores, it is possible to speculate about a rising intellectual differentiation.

Our study found that the secular changes in Estonia over the last 60 years were not on *g*, the general factor of intelligence. As such, they were not Jensen effects. Flynn (1999, 2000) has suggested that failures to find secular gains are due to the use of tests of "crystallized" intelligence because secular gains occur mostly in "fluid" intelligence. However, we agree with Jensen (1998, pp. 122–125) that the *g* factor makes the distinction between fluid and crystallized intelligence largely superfluous. For example, a prototypical test of crystallized intelligence is the Armed Services Vocational Aptitude Battery (ASVAB), which is used to select applicants for all military enlistments and assign them to first jobs. It consists of 10 separately scored subtests (General Science, Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Numerical Operations, Coding Speed, Auto and Shop Information, Mathematics Knowledge, Mechanical Comprehension, and Electronics Information). Yet, the *g* factor extracted from the ASVAB substantially correlates (.61) with speed of reaction time measures, a prototypical measure of fluid intelligence (for a review, see Jensen, 1998, pp. 234–238). This would not occur if fluid and crystallized intelligence were quite separate types. Flynn (personal communication, July 19, 2002) has responded to our empirical evidence by asserting that the distinction between Gf and Gc is indeed functional. Thus, much new data will be required to break this impasse and elucidate the full relation between the Jensen effect and the Flynn effect.

Acknowledgements

Supported by Grant 2387 from the Estonian Scientific Foundation. We thank all students and teachers for their participation in the study. We thank J. Philippe Rushton, Arthur R. Jensen, Ulric Neisser, and James R. Flynn for their helpful comments on an earlier version of this article.

References

Cattell, R. B. (1987). *Intelligence: its structure, growth, and action*. Amsterdam: Elsevier North-Holland.

Colom, R., Juan-Espinosa, M., & Garcia, L. F. (2001). The secular increase in test scores is a "Jensen effect". *Personality and Individual Differences*, *30*, 553–559.

Detterman, D. K., & Daniel, M. H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups. *Intelligence*, *13*, 349–359.

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

Flynn, J. R. (2000). IQ gains and fluid *g*. *American Psychologist*, *55*, 543.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf–Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (pp. 53–91). New York: Guilford Press.

Inglehart, R. (1995). Changing values, economic development and political change. *International Social Science Journal*, *145*, 379–404.

Jacobs, J. A. (1996). Gender inequality and higher education. *Annual Review of Sociology*, *22*, 153–185.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Jensen, A. R. (1998). *The g factor*. Westport, CT: Praeger.

Lynn, R. (1982). IQ in Japan and the United States shows a growing disparity. *Nature*, *297*, 222–223.

Neisser, U. (Ed.) (1998). *The rising curve: long term gains in IQ and related measures*. Washington, DC: American Psychological Association.

Raven, J. (2000). The Raven's Progressive Matrices: change and stability over culture and time. *Cognitive Psychology*, *41*, 1–48.

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

Rushton, J. P. (1999). Secular gains in IQ not related to the *g* factor and inbreeding depression—Unlike Black–White differences: a reply to Flynn. *Personality and Individual Differences*, *26*, 381–389.

Svensson, A., Emanuelsson, I., & Reuterberg, S. (1997, March 24–28). *Changes in intelligence from 1960 to 1995 in relation to cohort, gender, and socioeconomic background.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Terman, L. M. (1921). *The intelligence of school children*. London: George G. Harrap.

Tork, J. (1940). *Eesti laste intelligents (The intelligence of Estonian children)*. Tartu: Koolivara.