



PERGAMON

Personality and Individual Differences 27 (1999) 1221–1229

---

---

PERSONALITY AND  
INDIVIDUAL DIFFERENCES

---

---

www.elsevier.com/locate/paid

# Sex differences in historical knowledge and school grades: a 26 nation study

Sylvia Wilberg<sup>a</sup>, Richard Lynn<sup>b,\*</sup>

<sup>a</sup>*Phillips-University of Marburg, Germany*

<sup>b</sup>*University of Ulster, Coleraine, Northern Ireland, BT52 1SA, UK*

Received 21 August 1998

---

## Abstract

Sex differences among ninth grade school students in historical knowledge and grades were examined in 29 samples drawn from 26 countries. Boys obtained higher mean scores on the tests of historical knowledge in all samples, while girls obtained better grades in 22 of the samples. © 1999 Elsevier Science Ltd. All rights reserved.

*Keywords:* Sex differences; Gender differences; Historical knowledge; Educational achievement

---

## 1. Introduction

There is a large research literature on sex differences among adolescents in educational attainment. In general this has shown that boys tend to perform better than girls in physics, chemistry and mathematics, while girls tend to perform better than boys in language subjects including their own language and literature and foreign languages. For mathematics and the physical sciences, the largest recent data consist of the Third International Mathematics and Science Study (TIMSS) of 39 countries carried out in 1994/5. The results were that among 14 years olds boys performed better than girls in mathematics in 30 of the countries, of which 8 were statistically significant, while girls performed better than boys in 8 countries, none of which were statistically significant (Beaton, Martin, Mullis, Gonzales, Smith & Kelly, 1996a, p. 34). In physics and chemistry, boys performed better than girls in 38 countries, of which

---

\* Corresponding author. Tel: +44-1989-770908.

25 were statistically significant in physics and 16 statistically significant in chemistry (Beaton, Mullis, Gonzales, Smith & Kelly, 1996b, p.51).

In language subjects girls typically perform better than boys. For instance, in England approximately 90% of 15/16 year olds take the General Certificate of Secondary Education (GCSE). The results for 1993 were the 59% of girls of the total age group of approximately half a million achieved A–C grade passes in English as compared with 42% of the boys. In foreign languages, 37% of the total age group achieved A–C grade passes as compared with 24% of boys (Department of Education, 1994). In Ireland, it has been found that among 12–14 year olds who speak English as their first language but all of whom are required to learn Irish, girls perform slightly better than boys in English but substantially better in Irish (Lynn & Wilson, 1993).

In this paper we are concerned with sex differences among adolescents in history. This has been rather little investigated. In the English GCSE, girls perform better than boys. In 1993, 22% of girls of the total age group achieved A–C grade passes, as compared with 17% of boys (Department of Education, 1994). Similarly, among high school students in Germany and Austria, girls achieve better grades than boys by 0.15  $d$  ( $d$  = standard deviation units) and 0.21  $d$ , respectively (Hopp & Lienert, 1975; Weiss, 1976). However, in multiple-choice tests of historical knowledge it has been found in Germany that boys perform as well as girls (Ingenkamp, 1996; Furntratt, 1970). In the United States males perform better than females in history among 17/18 year olds taking the College Board Achievement Tests (Stumpf & Stanley, 1996), although these are not representative samples of school students.

The general trend of the existing research literature suggests that girls outperform boys in history grades, but this is not because girls have greater historical knowledge as tested in multiple-choice tests. In this paper we examine the issue of sex differences in knowledge and grades in history among ninth grade school students in 29 samples drawn from 26 countries.

## 2. Method

### 2.1. Samples

The samples consisted of 9th grade school students aged approximately 15 years drawn as representative samples from 26 European countries and Israel. In Israel, three samples were tested consisting of Arabs, Jews and Palestinians, and in Italy, two samples were tested, one from Rome and its environs and another from the South Tyrol. The study was carried out in 1994.

### 2.2. Measures

The information on grades was obtained by asking the students to report their last grade in History using a 5 point scale ranging from “very good” to “very bad”. Historical knowledge was assessed by 4 tests. Each of these consisted of five items which had to be placed in chronological sequence. These tests are given in the Appendix. Two of the tests were in verbal format and consisted of knowledge of economic and general history. The other two were in

pictorial format and consisted of knowledge of the history of costume and of ships. In designing the tests the investigators were concerned to examine the possibility that boys and girls might differ in the kinds of historical knowledge they possessed, such that boys might have more knowledge of traditionally male interests such as wars, weapons, machines, etc., while girls might have more knowledge of traditionally female interests. To investigate this possibility, one of the pictorial tests consisted of ships belonging to different historical periods, which it was considered might be of more interest to boys, and the other pictorial test consisted of couples dressed in the costumes of different historical periods, which it was considered might be of more interest to girls.

Table 1  
Boys' and girls' History test scores and grades in twenty-nine samples

Country	Sample size <i>N</i>	Test scores				Grades			
		Boys' mean (SD)	Girls' mean (SD)	<i>d</i>	<i>t</i>	Boys' mean (SD)	Girls' mean (SD)	<i>d</i>	<i>t</i>
Belgium	577	9.18 (3.19)	7.89 (2.52)	0.45	5.4 <sup>a</sup>	3.62 (0.93)	3.61 (0.95)	0.01	0.2
Bulgaria	1031	9.04 (3.74)	6.85 (3.60)	0.60	9.1 <sup>a</sup>	4.00 (0.96)	4.07 (0.96)	-0.07	1.1
Croatia	1025	7.76 (3.29)	6.99 (3.00)	0.24	3.9 <sup>a</sup>	3.14 (1.02)	3.48 (1.01)	-0.34	5.3 <sup>a</sup>
Czech Republic	1158	10.30 (2.62)	9.14 (2.58)	0.45	7.3 <sup>a</sup>	4.04 (0.79)	4.18 (0.79)	-0.19	3.2
Denmark	851	7.39 (3.28)	6.76 (2.89)	0.20	3.0	3.53 (0.89)	3.46 (0.79)	0.09	1.3
England	868	6.11 (3.37)	5.58 (3.32)	0.16	2.3	3.61 (0.92)	3.56 (0.85)	0.05	0.8
Estonia	1106	7.52 (2.75)	6.66 (2.95)	0.30	5.0 <sup>a</sup>	3.77 (0.75)	4.09 (0.71)	-0.44	7.2 <sup>a</sup>
Finland	899	7.55 (3.21)	7.05 (2.87)	0.16	2.4	3.38 (0.99)	3.73 (0.89)	-0.39	5.7 <sup>a</sup>
France	1001	8.70 (3.66)	8.20 (3.30)	0.14	2.3	3.36 (0.87)	3.30 (0.82)	0.07	1.0
Germany	2107	7.49 (3.08)	6.88 (2.84)	0.21	4.7 <sup>a</sup>	3.43 (0.75)	3.40 (0.73)	0.04	0.9
Greece	1724	6.89 (3.27)	6.07 (2.96)	0.26	5.5 <sup>a</sup>	3.80 (0.96)	3.94 (0.97)	-0.14	2.9
Hungary	941	8.63 (3.17)	7.75 (2.93)	0.29	4.4 <sup>a</sup>	3.36 (0.90)	3.61 (0.83)	-0.28	4.3 <sup>a</sup>
Iceland	967	7.29 (3.66)	6.10 (2.95)	0.36	5.6 <sup>a</sup>	3.72 (0.92)	3.81 (0.91)	-0.10	1.5
Israel—Arabs	356	4.18 (3.14)	3.18 (2.58)	0.35	3.2 <sup>a</sup>	4.03 (1.09)	4.20 (0.86)	-0.18	1.7
Israel—Jews	1144	6.36 (3.45)	5.38 (3.24)	0.29	4.9 <sup>a</sup>	3.88 (0.94)	4.04 (0.86)	-0.17	2.9
Israel—Palestines	1099	3.98 (3.03)	3.87 (3.20)	0.01	0.1	3.93 (1.07)	4.16 (0.92)	-0.23	3.7 <sup>a</sup>
Italy	1256	8.77 (3.53)	8.39 (3.45)	0.11	1.9	3.18 (0.82)	3.42 (0.82)	-0.29	5.0 <sup>a</sup>
Italy—South Tyrol	812	8.23 (2.97)	7.05 (2.78)	0.41	5.9 <sup>a</sup>	3.41 (0.85)	3.52 (0.81)	-0.13	1.9
Lithuania	1216	8.82 (2.79)	7.22 (2.91)	0.56	9.5 <sup>a</sup>	3.81 (0.60)	3.90 (0.63)	-0.15	2.5
Norway	1237	7.57 (3.51)	6.61 (3.19)	0.29	5.0 <sup>a</sup>	3.60 (0.94)	3.70 (0.82)	-0.11	1.9
Poland	1280	7.04 (2.93)	6.18 (2.85)	0.30	5.3 <sup>a</sup>	3.79 (0.94)	4.02 (0.88)	-0.25	4.4 <sup>a</sup>
Portugal	1254	6.16 (3.29)	5.13 (3.16)	0.32	5.6 <sup>a</sup>	3.24 (0.92)	3.25 (0.95)	-0.01	0.2
Russia	1785	10.13 (3.52)	9.33 (3.22)	0.24	4.9 <sup>a</sup>	3.87 (0.73)	3.98 (0.68)	-0.16	3.4 <sup>a</sup>
Scotland	242	6.56 (3.35)	5.82 (3.37)	0.22	1.7	3.64 (0.92)	3.79 (0.91)	-0.16	1.2
Slovenia	883	6.31 (3.50)	5.25 (3.17)	0.32	4.7 <sup>a</sup>	3.63 (1.10)	3.65 (1.04)	-0.02	0.2
Spain	991	8.20 (3.04)	7.36 (2.86)	0.28	4.5 <sup>a</sup>	3.60 (0.95)	3.55 (0.90)	0.06	0.9
Sweden	829	7.64 (3.55)	6.94 (3.25)	0.21	3.0	3.46 (0.91)	3.55 (0.87)	-0.11	1.5
Turkey	1229	5.23 (3.03)	4.99 (2.84)	0.08	1.4	3.87 (0.83)	3.85 (0.88)	0.03	0.4
Ukraine	1148	8.37 (3.29)	7.59 (3.62)	0.23	3.8 <sup>a</sup>	4.04 (0.76)	4.29 (0.74)	-0.33	5.5 <sup>a</sup>

<sup>a</sup>  $p < 0.05$  (two-tailed *t*-test) after a correction for multiple testing [ $1-(1-\alpha)^k = 0.05$ , with  $k = 29$ ]; positive *ds* denote higher means of boys, negative *ds* denote higher means of girls.

The tests were scored as follows: if the historical sequence was fully correct, 4 points were awarded; if one item was incorrectly placed, 2 points were awarded; if three items were wrongly paced but both the first and last item were correct, 1 point was awarded. Thus, the possible scores for each task ranged between 0–4, and the possible scores for the whole test ranged between 0–16.

### 3. Results

The principal results are shown in Table 1. This gives for each country the sample size, the means and standard deviations of the test scores on the historical knowledge test, and the *ds* (the differences between the means for boys and girls divided by the pooled standard deviation weighting the sds for boys and girls equally). The statistical significance of the differences in the scores of boys and girls was tested by *t*-tests with Bonferroni corrections for multiple testing. The same information is given for grades in the right hand section of the table.

The results show that on the historical knowledge test boys obtained higher means than girls in all the samples and that in 20 of these the differences are statistically significant. The mean unweighted advantage of boys was 0.28 *d* (0.28 of a standard deviation).

In contrast, girls obtained better grades in 22 of the samples, of which 10 are statistically significant, while boys reported better grades in 7 of the samples of which none are statistically significant. The mean unweighted advantage of the girls was 0.13 *d* (0.13 of a standard deviation). Thus, the general conclusion to be drawn from these results is that boys tend to score higher on historical knowledge while girls tend to obtain better grades.

To examine whether there was any relationship between the test scores of historical knowledge and grades, correlations were calculated for boys and girls separately. The correlations were all positive and ranged from 0.09 to 0.36. The median correlations were 0.28 for boys and 0.22 for girls. These are highly statistically significant on such large sample sizes but are also quite low, indicating that the amount of historical knowledge is not strongly reflected in grades.

We turn now to the subsidiary objective of the investigation, concerned with the possibility that boys and girls might have different kinds of historical knowledge, and that boys might have more knowledge of ships while girls might have more knowledge of costumes. The results

Table 2

Mean scores of boys and girls (nb = 14,787, ng = 16,229) in the four components of the historical knowledge test

Test	Boys		Girls		<i>d</i>	<i>t</i>
	Mean	SD	Mean	SD		
Economic	2.64	1.64	2.59	1.62	0.03	2.5 <sup>a</sup>
General	0.43	1.10	0.30	0.93	0.13	10.9 <sup>a</sup>
Costume	2.55	1.64	2.41	1.64	0.09	7.3 <sup>a</sup>
Ships	1.91	1.12	1.50	1.06	0.39	33.4 <sup>a</sup>

<sup>a</sup>  $p < 0.01$ .

for the entire sample are given in Table 2. It will be seen that the boys have a slight but statistically significant advantage on the questions on economic history, general history and costume history, with  $d$ -values of 0.03, 0.13 and 0.09, respectively. There is little difference between the knowledge of boys and girls in these three areas. Thus, the hypothesis that girls would have a better knowledge of costume history because costume is a feminine interest has to be regarded as disconfirmed. However, the boys have a considerably greater advantage on the ships test (0.19  $d$ ) than in the other three areas. This provides some support for the hypothesis that boys and girls have different kinds of historical knowledge.

## 5. Discussion

The results contain three principal points of interest. First, boys obtained higher mean scores than girls on the test of historical knowledge in all 29 samples and in 20 of which this advantage is statistically significant. The boys' advantage is present and statistically significant in all of the four areas of historical knowledge tested, but it is substantially greater on the ships question than on the economic, general and costume history questions. This indicates that there are sex differences in the kinds of historical knowledge. These probably reflect the interests of boys and girls. Boys are likely to be more interested in machines of various kinds and ships are part of this general interest. However, the hypothesis that girls would be more interested in costume and have more historical knowledge of it was not confirmed. It remains an interesting question whether it would be possible to find an area of historical knowledge in which girls have greater interest and knowledge than boys.

Second, girls obtained higher grades than boys in 22 of the samples, of which 10 are statistically significant. The general trend of the results on grades is clearly that girls tend to get better grades than boys although this is less pronounced than the tendency of boys to have greater historical knowledge.

Third, the general pattern of results indicating that boys tend to do better on tests while girls tend to do better on grades is consistent with previous research. For instance, Kimball (1989) found this pattern for mathematics in the United States. Other American studies have found that in universities women obtain better grades than would be predicted from their Scholastic Aptitude Test (SAT) scores, as compared with men who obtain lower scores than would be predicted (Linn, 1990; Young, 1991, 1994).

There are two probable explanations for this sex difference in grades. The first is that females tend to work more conscientiously than males. Grades are typically awarded largely on the basis of course work and reflect the amount of effort devoted to the work on the basis of which grades are awarded. Several studies have found that females tend to have a stronger work ethic than males and this is likely to be expressed in more conscientious course work, e.g., Farmer (1983); Lynn (1991); Nevill and Perrotta (1988); Nevill and Super (1988); Luzzo (1994); Watson and Steed (1990). The second probable reason why females tend to obtain higher grades than males is that they tend to have better language abilities including essay writing skills, vocabulary and word fluency (Hyde & Linn, 1988) and this also is likely to contribute to better course work.

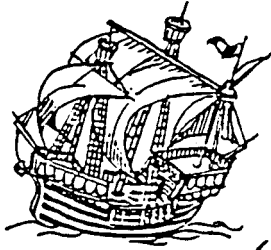
**Appendix A**

In these two questions, place the events in order in which they happened in the history of the world, from the earliest (1) to the most recent (5). Put one number in each bracket.

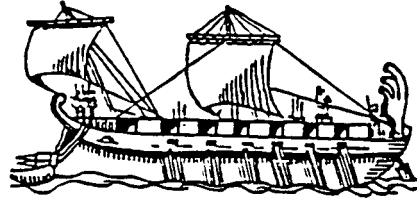
- a. Waged work in factories
- b. Trade between towns
- c. The cultivation of land
- d. Hunting and gathering
- e. Automatic mass production

- a. The Russian Revolution
- b. The Second World War
- c. The decolonisation of Africa
- d. The foundation of the UN
- e. The great world depression (the Wall Street Crash)

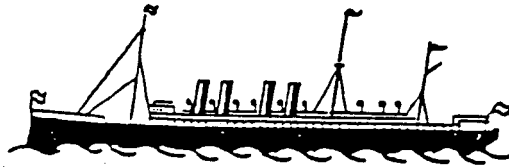
Put the following ships into the correct chronological order from the oldest (1) to the newest (5).



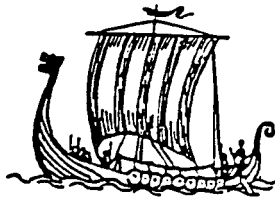
( )



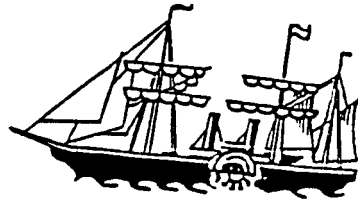
( )



( )

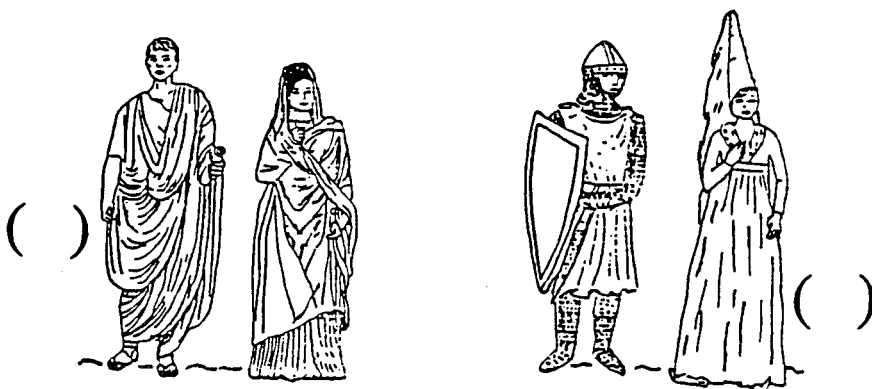


( )



( )

Put the following couples into the correct chronological order from the oldest (1) to the newest (5).





## References

- Beaton, A. E., Martin, M. O., Mullis, I. V., Gonzales, E. J., Smith, T. A., & Kelly, D. L. (1996a). *Mathematics achievement in the middle school years*. MA, TIMSS: Chestnut Hill.
- Beaton, A. E., Martin, M. O., Mullis, I. V., Gonzales, E. J., Smith, T. A., & Kelly, D. L. (1996b). *Science achievement in the middle school years*. MA, TIMSS: Chestnut Hill.
- Department of Education (1994). In *GCSE and GCE A/AS examination results 1992/3*, 7. Statistical Bulletin.
- Farmer, H. S. (1983). Career and homemaking plans for high school youth. *Journal of Counselling Psychology*, 30, 40–45.
- Furntratt, E. (1970). *Differentieller-Wissens-Test DWT, Handanweisung*. Göttingen: Hogrefe.
- Hopp, A. D., & Leinert, G. (1975). Eine Verteilungsanalyse von Gymnasialzensuren. *Schule und Psychologie*, 5, 139–150.
- Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: a meta-analysis. *Psychological Bulletin*, 104, 53–69.
- Ingenkamp, K. (1996). *Geschichtstest Neuzzeit GTN 8-10 Teil: 1890–1932*. Weinheim: Beltz.
- Kimball, M. (1989). A new perspective on women's math achievement. *Psychological Bulletin*, 105, 198–214.
- Linn, R. L. (1990). Admissions testing: recommended uses, validity, differential prediction and coaching. *Applied Measurement in Education*, 3, 297–318.
- Luzzo, D. A. (1994). An analysis of gender and ethnic differences in college students' commitment to work. *Journal of Employment Counselling*, 31, 38–45.
- Lynn, R. (1991). *The secret of the Miracle Economy*. London Social Affairs Unit.
- Lynn, R., & Wilson, R. G. (1993). Sex differences in cognitive abilities among Irish primary and secondary school children. *Irish Journal of Psychology*, 14, 293–300.
- Nevill, D. D., & Perrotta, J. M. (1988). Adolescents perceptions of work and home in Australia, Portugal and the U.S.A. *Journal of Cross-Cultural Psychology*, 16, 483–495.
- Nevill, D. D., & Super, D. E. (1988). Career maturity and commitment to work in university students. *Journal of Vocational Behavior*, 32, 139–151.
- Stumpf, H., & Stanley (1996). Gender-related differences on the college board's advanced placement and achievement tests, 1982–1992. *Journal of Educational Psychology*, 88, 363–364.
- Watson, M. B., & Steed, G. B. (1990). Work-role salience of South African adolescents. *Journal of Vocational Behavior*, 36, 249–257.
- Weiss, R. (1976). Über die Strenge der Benotung in verschiedenen Unterrichtsgegenständen. In K. Ingenkamp, *Die Fragwürdigkeit der Zensurengebung*. Weinheim: Beltz.
- Young, J. W. (1991). Gender bias in predicting college grades by gender and ethnicity: a new approach using item response theory. *Journal of Educational Measurement*, 28, 37–47.
- Young, J. W. (1994). Differential prediction of college grades by gender and ethnicity: a replication study. *Educational and Psychological Measurement*, 54, 1022–1029.