Personality and Individual Differences 51 (2011) 321–324

Contents lists available at ScienceDirect

Personality and Individual Differences

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

A longitudinal study of sex differences in intelligence at ages 7, 11 and 16 years

Richard Lynn [a,*], Satoshi Kanazawa [b]

[a] University of Ulster, Coleraine, Northern Ireland, BT52 1SA, UK
[b] London School of Economics and Political Science, Houghton St., London WC2A 2AE, UK

ARTICLE INFO

Article history:
Received 16 November 2010
Received in revised form 23 February 2011
Accepted 24 February 2011
Available online 7 May 2011

Keywords:
IQ
Sex differences
Maturation
Age trends

ABSTRACT

This paper presents the results of a longitudinal study of sex differences in intelligence as a test of Lynn's (1994) hypothesis that from the age of 16 years males develop higher average intelligence than females. The results show that at the ages of 7 and 11 years girls have an IQ advantage of approximately 1 IQ point, but at the age of 16 years this changes in the same boys and girls to an IQ advantage of 1.8 IQ points for boys.

1. Introduction

From the early years of the twentieth century it has been consistently asserted that there is no sex difference in average general intelligence defined as the sum of cognitive abilities measured by the IQ obtained in tests like the Wechslers and the Binets. In an early study, the absence of a sex difference in average intelligence was asserted by Terman (1916) who wrote that in the American standardization sample of the Stanford–Binet test on 4–16 year olds "the superiority of girls over boys is so slight . . . that for practical purposes it would seem negligible". The same view was taken by Burt and Moore (1912) and Spearman (1923). In the second half of the century it was reaffirmed by Cattell (1971, p. 131): "it is now demonstrated by countless and large samples that on the two main general cognitive abilities – fluid and crystallized intelligence – men and women, boys and girls, show no significant differences"; Hutt (1972, p. 88): "there is little evidence that men and women differ in average intelligence"; Maccoby and Jacklin (1974, p. 65): "the sexes do not differ consistently in tests of total (or composite) abilities"; Jensen (1980, p. 360): "males and females do not differ in IQ"; Brody (1992, p. 323): "gender differences in general intelligence are small and virtually non-existent"; and Herrnstein and Murray (1994, p. 275): "the consistent story has been that men and women have nearly identical IQs".

This consensus was broken by Lynn (1994, 1999) who contended that while it is correct that there is virtually no sex difference in average intelligence between the ages of 5 and 15 years, from the age of 16 years males begin to have greater average intelligence than females and that this increases to an advantage of between 4 and 5 IQ points in adults. More specifically, Lynn (1994) proposed that there is virtually no sex difference in intelligence between the ages of 5–10 years, that between the ages of 11–14 years girls have a small IQ advantage of approximately 1 IQ point because they mature earlier, and that from the age of 15–16 years boys develop a small IQ advantage of approximately 1 IQ point, which increases in later adolescence to reach approximately 4 IQ points among adults.

Lynn's (1994) hypothesis was first disputed by Mackintosh (1996) on the grounds that a review by Court (1983) had shown that there is no sex difference on the Progressive Matrices, from which Mackintosh (1996, p. 567) concluded that "there is no sex difference in general intelligence worth speaking of". In response to this, Lynn and Irwing (2004) published a meta-analysis of 57 studies of sex differences on the Progressive Matrices in which they showed that there is no difference among children aged 6–14 years, but that males obtain higher means from the age of 15 through to old age, and that among adults, the male advantage is 5 IQ points. A year later Irwing and Lynn (2005) published a meta-analysis of 22 studies of sex differences on the Progressive Matrices in university students and concluded that in these samples males have an advantage of 4.6 IQ points. In a more recent study, Mackintosh and Bennett (2005) reported data for a sample of 17 year olds on the Progressive Matrices in which males obtained a higher mean of 6.4 IQ points. In this paper they conceded that "studies of older participants (over the age of 16) were more likely to yield a male than a female advantage" (p. 670).

Lynn's hypothesis has been confirmed in the Spanish standardization sample of the WAIS–III, in which men obtained a higher IQ than women of 3.6 IQ points (Colom, Garcia, Juan-Espinoza, &

* Corresponding author.
E-mail address: lynnr540@aol.com (R. Lynn).

0191-8869/$ - see front matter
doi:10.1016/j.paid.2011.02.028

Abad, 2002). A further confirmation for a Spanish sample has been reported by Colom and Lynn (2004), who found a male advantage among 18 year olds of 4.3 IQ points on the Differential Aptitude Test. Further supportive evidence for Lynn's hypothesis has been published by Meisenberg (2009), who reports a male advantage of 2.8 IQ points among 22–3 year old whites in the United States on the ASVAB (Armed Services Vocational Aptitude Battery). This difference, however, was not present among blacks. In this study intelligence was also measured as *g*, and for this there was no significant sex difference among 15 year olds among either blacks or whites, but among whites a significant male advantage of 4 IQ points was present among 16 year olds and this increased to an advantage of 6.5 among 22–3 year olds, while for blacks a male advantage of 1 IQ point was present among 16 year olds and this increased to an advantage of 2.15 points among 22–3 year olds. Three studies in Germany have reached the same conclusion for the IST test. Amelang and Steinmayr (2006) report a male advantage of 6.0 IQ points in a sample of 34 year olds. Steinmayr and Spinath (2010) report a male advantage of 9.7 IQ points in a sample of 16 year olds; and Steinmayr, Beauducel, and Spinath (2010) report a male advantage of 11.5 IQ points in a further sample of 16 year olds.

Lynn's hypothesis has been further confirmed in a review by Ellis et al. (2008, p. 288) of sex differences in general intelligence defined as the IQ obtained in tests like the Wechsler. He lists 50 studies of adults. Males obtained statistically significantly higher IQs than females in 29 studies and there was no statistically significant difference in 20 studies. In evaluating the non-statistically significant studies, it should be borne in mind that a sample size of around 500 is required to obtain a statistically significantly difference of 5 IQ points and many of the studies fall short of this number. There was one study in which females obtained higher IQs than males, but this was of a mentally subnormal sample and should be discounted because males are more impaired in mentally subnormal samples (Ellis et al., 2008, p. 290). Thus, the preponderance of the evidence reviewed by Ellis et al. (2008) indicates that Lynn's hypothesis that men have a higher average IQ than women is correct.

Despite these results supporting Lynn's hypothesis, many scholars continue to assert that there is no average sex difference in intelligence. For instance, Halpern (2000, p. 218): "sex differences have not been found in general intelligence"; Butterworth (1999, p. 293): "women's brains are 10% smaller than men's, but their IQ is on average the same"; Geary (1998, p. 310): "the overall pattern suggests that there are no sex differences, or only a very small and unimportant advantage of boys and men, in average IQ scores"; Bartholomew (2004, p. 91): "men on average have larger brains than women but display no significant advantage in cognitive performance"; Anderson (2004, p. 829): "it is an important finding of intelligence testing that there is no difference between the sexes in average intellectual ability; this is true whether general ability is defined as an IQ score calculated from an omnibus test of intellectual abilities such as the various Wechsler tests, or whether it is defined as a score on a single test of general intelligence, such as Raven's Matrices"; Hines (2007, p. 103) "there appears to be no sex difference in general intelligence; claims that men are more intelligent than women are not supported by experimental data; Haier (2007, p. 115): "general intelligence does not differ between men and women"; Halpern (2007, p. 123) "there is no difference in intelligence between males and females…overall, the sexes are equally smart"; Speke (2007, p. 65): "men and women have equal cognitive capacity".

It is apparent that Lynn's hypothesis has not been widely accepted. Our objective in this paper is to present a test of the hypothesis using longitudinal data rather than the cross-sectional data on which the hypothesis has hitherto been based. Cross-sectional data are problematic because among older adolescents, more males are in prison and other custodial institutions and these are not included in the samples on which norms are based. These excluded males have lower average IQs, so their omission from normative samples inflates the IQs of males that are tested. Longitudinal data overcome this problem by assessing IQs of the same males and females at different ages and hence provide a more stringent test of the hypothesis.

2. Method

The data to be analysed come from the National Child Development Study (NCDS). NCDS is a large-scale prospectively longitudinal study which has followed a population of British respondents since birth for more than half a century. The study began with all babies (n = 17,419) born in Great Britain (England, Wales, and Scotland) during the week of March 03–09, 1958. The respondents are subsequently re-interviewed in 1965 (Sweep 1 at age 7; n = 15,496), in 1969 (Sweep 2 at age 11; n = 18,285) (There are more respondents in Sweep 2 than in the original sample (Sweep 0) because the Sweep 2 sample includes eligible children who were in the country in 1969 but not in 1958 when Sweep 0 interviews were conducted), in 1974 (Sweep 3 at age 16; n = 14,469), in 1981 (Sweep 4 at age 23; n = 12,537), in 1991 (Sweep 5 at age 33; n = 11,469), in 1999–2000 (Sweep 6 at age 41–42; n = 11,419), and in 2004–2005 (Sweep 7 at age 46–47; n = 9534). In each Sweep, personal interviews and questionnaires were administered to the respondents, to their mothers, teachers, and doctors during childhood and to their partners and children in adulthood. 97.8% of the NCDS respondents are Caucasian.

The NCDS respondents took multiple intelligence tests at ages 7, 11, and 16. At age 7, the respondents took four cognitive tests (Copying Designs Test, Draw-a-Man, Southgate Group Reading, and Problem Arithmetic). At age 11, they took five cognitive tests (Verbal General Ability, Nonverbal General Ability, Reading Comprehension, Mathematics, and Copying Designs). At age 16, they took two cognitive tests (Reading Comprehension and Mathematics Comprehension). We first perform a factor analysis at each age to compute their general intelligence score for each age. All cognitive test scores at each age load on only one latent factor, with reasonably high factor loadings (Age 7: Copying Designs = .671, Draw-a-Man = .696, Southgate Group Reading = .780, and Problem Arithmetic = .762; Age 11: Verbal General Ability = .920, Nonverbal General Ability = .885, Reading Comprehension = .864, Mathematics = .903, and Copying Designs = .486; Age 16: Reading Comprehension = .909, and Mathematics Comprehension = .909). The latent general intelligence factors at each age are converted into the standard IQ metric, with a mean of 100 and a standard deviation of 15.

3. Results

Table 1 gives descriptive statistics for test scores of the boys and girls at ages 7, 11 and 16 years. Table 2 shows the mean IQs and sds of all the participants tested at the three ages. The two columns at the right of the table give the sex differences expressed in standard deviation units (d), and the values of Student's t as tests of the statistical significance of the differences. It will be seen that at the ages of 7 and 11 girls obtained a higher average IQ than boys, but at the age of 16 years boys obtained a higher average IQ than girls. All the sex differences are statistically significant at $p < .01$.

It will be noted that the numbers tested decline over the three age groups. By age 16 years, 16.9% of the original sample at age 0 (and 6.6% of the age 7 sample) had dropped out. This leaves open the possibility of greater attrition among lower IQ boys. To exam-

Table 1
Descriptive statistics for test scores at ages 7,11 and 16 years.

	Boys	Girls
Age 7		
Copying Designs	7.06	6.97
Draw-a-man	23.56	24.14
Southgate group reading	22.44	24.29
Problem arithmetic	5.22	5.00
Age 11		
Verbal general ability	21.05	23.13
Nonverbal general ability	20.76	21.01
Reading Comprehension	15.93	16.03
Mathematical	16.81	16.44
Copying Designs	8.38	8.30
Age 16		
Reading Comprehension	25.35	25.27
Mathematics comprehension	13.39	12.09

Table 2
IQs of all tested boys and girls aged 7, 11 and 16 years. Negative signs denote higher means obtained by girls.

Age	Boys			Girls			Sex difference	
	Mean	SD	*n*	Mean	SD	*n*	*d*	*t*
7	99.447	15.225	7401	100.585	14.737	7006	−.08	−4.555
11	99.596	15.334	7240	100.427	14.628	6855	−.06	−3.288
16	100.780	15.627	6103	99.182	14.269	5816	.11	5.822

Table 3
IQs of the same boys and girls tested at the ages of 7, 11 and 16 years. Negative signs denote higher means obtained by girls.

Age	Boys			Girls			Sex difference	
	Mean	SD	*n*	Mean	SD	*n*	*d*	*t*
7	100.374	14.693	4626	101.420	14.262	4458	−.07	−3.443
11	100.678	14.968	4626	101.520	14.216	4458	−.06	−2.747
16	101.461	15.253	4626	99.681	14.085	4458	.12	5.775

Table 4
Factor loadings for the three separate multi-group confirmatory factor analysis.

	Boys	Girls
Age 7		
Copying Designs test	.665	.678
Draw-a-man test	.693	.695
Southgate group reading test	.786	.783
Problem arithmetic test	.768	.763
Age 11		
Verbal general ability test	.924	.922
Nonverbal general ability test	.884	.885
Reading Comprehension test	.864	.863
Mathematical test	.909	.899
Copying Designs test	.477	.500
Age 16		
Reading Comprehension test	.913	.906
Mathematics comprehension test	.913	.906

Table 5
Factor loadings in confirmatory factor analysis and effect size of the sex differences in test scores.

	Factor loading	Sex difference in Cohen's *d*
Age 7		
Copying Designs test	.671	.045
Draw-a-man test	.696	−.082
Southgate group reading test	.780	−.259
Problem arithmetic test	.762	.088
Age 11		
Verbal general ability test	.920	−.222
Nonverbal general ability test	.885	−.033
Reading Comprehension test	.864	−−.016
Mathematical test	.903	.036
Copying Designs test	.486	.053
Age 16		
Reading Comprehension test	.909	.011
Mathematics comprehension test	.909	.186

ine this possibility, Table 3 shows the mean IQs and sds of only the participants who took the IQ tests at all three ages. It will be seen that the results are virtually identical to those given in Table 2. All the sex differences are statistically significant at $p < .01$.

We have carried out a multi-group confirmatory factor analysis with mean structures in order to test for measurement invariance using the procedure described by Wicherts and Dolan (2010). We first performed a multi-group confirmatory factor analysis (in this case, with two groups – boys and girls). Table 4 presents the factor loadings for the three separate multi-group confirmatory factor analyses at the three different ages. It will be seen that the factor loadings for boys and for girls for all the tests are virtually identical. This satisfies the first condition of establishing measurement invariance proposed by Wicherts and Dolan (2010).

To examine this issue further, we look now at the factor loadings in confirmatory factor analysis and at whether the effect sizes of the sex differences in the test scores are collinear. The results are shown in Table 5. It will be seen that the factor loadings and the absolute values of the sex differences expressed as *d* are largely collinear. According to Wicherts and Dolan, this establishes the method invariance and the equivalence of both slopes and intercepts. The Pearsonian product-moment correlation between the factor loading and the absolute value of the sex difference expressed as *d* is $r = .779$ at age 7 and $r = .215$ at age 11. There are not sufficient data points to compute a bivariate correlation at age 16. Having only two indicators at age 16 also forces us to have the identical factor loading for both of them. The Spearman rank-order correlation is $\rho = 1.000$ at age 7 and $\rho = .400$ at age 11.

It should be noted that the IQ scores computed at all three ages are simple linear transformations of the latent factor extracted from the factor analysis: IQ = 100 + 15 × (latent factor), i.e. the latent factor was multiplied by 15. Thus, we are comparing the sexes on a linear transformation of the latent factor.

4. Discussion

The principal interest of the results is that these longitudinal data show that the same girls, who obtained a higher average IQ than boys at the ages of 7 and 11 years, obtained a lower average IQ than boys at the age 16 years. This result confirms Lynn's hypothesis that the intelligence of boys and girls matures at different rates, and that the earlier maturation of girls gives them an IQ advantage at the ages of 7 and 11 years, while the later maturation of boys gives them an IQ advantage at the age of 16 years. This result disconfirms the hypothesis advanced by Madhyastha, Hunt, Deary, and Dykiert (2009) that the mean intelligence of females declines relative to males' over time in longitudinal surveys because of differential attrition, where low-IQ men are more likely to drop out of the surveys than low-IQ women. To test this proposition, it is shown that the reversal in sex difference in intelligence before and after puberty is found when the same sample is tested at different ages, and therefore is not the result of differential attrition of low-IQ men and women.

The IQ advantage of boys at the age 16 years in this data set is .12*d* and is equivalent to 1.8 IQ points. This is virtually the same as the male advantage proposed at this age by Lynn (1994) and

Lynn and Irwing (2004). As noted in the introduction, the confirmation of Lynn's hypothesis in these longitudinal data provides a more stringent test of Lynn's hypothesis than has been provided by cross-sectional data. While the present results show that in later adolescence males surpass females in intelligence in this British sample, this finding will not necessarily be replicated in other populations. Hopefully, this article will challenge other researchers to conduct similar studies elsewhere in the world. It may also be of interest to note that boys had larger standard deviations than girls at all three ages, confirming the frequent contention advanced by Havelock Ellis (1904) and many others that the variance of intelligence is greater in males than in females.

It is an interesting question why girls mature earlier than boys. One reason may be in human evolutionary history of mild polygyny. Under polygyny, girls who reach puberty earlier gain a reproductive advantage over their age mates by being able to marry polygynous men. The reason for this is that females are in competition *with* other females for reproductive success. Females who mature earlier than other females can start reproducing earlier and have access to polygynous men, while their age mates who have not yet reached puberty cannot. At the same time, there is no reproductive incentive for men in a polygynous breeding system to mature earlier. Another possible reason for the earlier maturation of girls than boys is that decline in gamete quality with age is a more serious constraint for women than for men. Decline in sperm quality is buffered by the large number of sperm produced. In contrast, ova production is much more limited. Women are therefore under stronger selection pressure to begin mating as soon as possible.

References

Amelang, M., & Steinmayr, R. (2006). Is there a validity increment for tests of emotional intelligence in explaining the variance of performance criteria? *Intelligence, 34*, 459–468.

Anderson, M. (2004). Sex differences in general intelligence. In R. L. Gregory (Ed.), *The Oxford companion to the mind*. Oxford, UK: Oxford University Press.

Bartholomew, D. J. (2004). *Measuring intelligence. Facts and fallacies*. Cambridge: Cambridge University Press.

Brody, N. (1992). *Intelligence*. San Diego, CA: Academic.

Burt, C. L., & Moore, R. C. (1912). The mental differences between the sexes. *Journal of Experimental Pedagogy, 1*, 355–388.

Butterworth, B. (1999). *The Mathematical Brain*. London: Macmillan.

Cattell, R. B. (1971). *Abilities: Their structure, growth and action*. Boston: Houghton Mifflin.

Colom, R., Garcia, L. F., Juan-Espinoza, M., & Abad, F. J. (2002). Null sex differences in general intelligence. Evidence from the WAIS–III. *Spanish Journal of Psychology, 5*, 29–35.

Colom, R., & Lynn, R. (2004). Testing the developmental hypothesis of sex differences in intelligence on 12–18 year olds. *Personality and Individual Differences, 36*, 75–82.

Court, J. H. (1983). Sex differences in performance on Raven's Progressive Matrices: A review. *Alberta Journal of Educational Research, 29*, 54–74.

Ellis, H. (1904). *Man and woman: A study of human secondary sexual characteristics*. London: Walter Scott.

Ellis, L., Hershberger, S., Field, E., Wersinger, S., Pellis, S., Geary, D., et al. (2008). *Sex differences: Summarizing more than a century of scientific research*. New York: Psychology Press.

Geary, D. C. (1998). *Male, female*. Washington, DC: American Psychological Association.

Haier, R. (2007). Brains, bias, and biology: Follow the data. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't there more women in science?* Washington, DC: American Psychological Association.

Halpern, D. (2000). *Sex differences in cognitive abilities*. Mahwah, NJ: Lawrence Erlbaum.

Halpern, D. (2007). Science, sex and good sense: Why women are underrepresented in some areas of science and math. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't there more women in science?* Washington, DC: American Psychological Association.

Herrnstein, R., & Murray, C. (1994). *The bell curve*. New York: Random House.

Hines, M. (2007). Do sex differences in cognition cause the shortage of women in science? In S. J. Ceci & W. M. Williams (Eds.), *Why aren't there more women in science?* Washington, DC: American Psychological Association.

Hutt, C. (1972). *Males and females*. Harmondsworth, UK: Penguin Books.

Irwing, P., & Lynn, R. (2005). Sex differences in means and variability on the Progressive Matrices in university students: A meta-analysis. *British Journal of Psychology, 96*, 505–524.

Jensen, A. R. (1980). *Bias in mental testing*. London: Methuen.

Lynn, R. (1994). Sex differences in brain size and intelligence. A paradox resolved. *Personality and Individual Differences, 17*, 257–271.

Lynn, R. (1999). Sex differences in intelligence and brain size: A developmental hypothesis. *Intelligence, 27*, 1–12.

Lynn, R., & Irwing, P. (2004). Sex differences on the Progressive Matrices: A meta-analysis. *Intelligence, 32*, 481–498.

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Mackintosh, J. M., & Bennett, E. S. (2005). What do Raven's Matrices measure? An analysis in terms of sex differences. *Intelligence, 33*, 663–674.

Mackintosh, N. J. (1996). Sex differences and IQ. *Journal of Biosocial Science, 28*, 559–572.

Madhyastha, T. M., Hunt, E., Deary, I. J., & Dykiert, D. (2009). Recruitment modeling applied to longitudinal studies of group differences in intelligence. *Intelligence, 37*, 422–4427.

Meisenberg, G. (2009). Intellectual growth during late adolescence. Sex and race. *Mankind Quarterly, 50*, 138–155.

Spearman, C. (1923). *The nature of intelligence and the principles of cognition*. London: Macmillan.

Speke, E. (2007). Sex, math and science. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't there more women in science?* Washington, DC: American Psychological Association.

Steinmayr, R., Beauducel, A., & Spinath, B. (2010). Do sex difference in a faceted model of fluid and crystallized intelligence depend on the method applied? *Intelligence, 38*, 101–110.

Steinmayr, R., & Spinath, B. (2010). Sex differences in school achievement. *Journal of Personality, 22*, 185–209.

Terman, L. M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.

Wicherts, J. M., & Dolan, C. V. (2010). Measurement invariance in confirmatory factor analysis: An illustration using IQ test performance of minorities. *Educational Measurement: Issues and Practice, 29*, 39–47.