Available online at www.sciencedirect.com
SCIENCE DIRECT®

Personality and Individual Differences 39 (2005) 103–114

PERSONALITY AND INDIVIDUAL DIFFERENCES

www.elsevier.com/locate/paid

Sex differences on the WISC-R in New Zealand

Richard Lynn [a,*], David M. Fergusson [b], L. John Horwood [b]

[a] *University of Ulster, Coleraine, Northern Ireland, UK*
[b] *Christchurch Health and Development Study, Christchurch School of Medicine & Health Sciences, PO Box 4345, Christchurch, New Zealand*

Received 15 July 2004; received in revised form 15 November 2004; accepted 13 December 2004
Available online 26 February 2005

Abstract

Sex differences on the WISC-R were examined in a sample of 897 New Zealand children studied at ages 8 and 9 years. Boys scored significantly higher than girls on the subtests of information, vocabulary, block design and object assembly, while girls scored significantly higher on coding. Boys obtained slightly but not significantly higher scores on the verbal, performance and full scale IQs. The results were in general similar to the sex differences in the standardisation samples of the WISC-R in Scotland, the Netherlands and the United States.

Keywords: Sex differences; WISC-R; Intelligence

1. Introduction

The question of whether there are sex differences in intelligence has probably been debated from time immemorial and certainly from the early twentieth century following the construction of intelligence tests and collection of IQs of large samples of males and females. It has virtually invariably been considered that there is no sex difference in general intelligence. This conclusion was reached in the second decade of the twentieth century by Terman (1916, pp. 69–70) on the

* Corresponding author. Tel.: +44 1275 392 092.
E-mail address: lynnr540@aol.com (R. Lynn).

0191-8869/$ - see front matter
doi:10.1016/j.paid.2004.12.009

basis of his American standardisation sample of the Stanford–Binet test on approximately 1000 4–16 year olds. In this sample, girls obtained a slightly higher average IQ than boys but "the superiority of girls over boys is so slight... that for practical purposes it would seem negligible". It has, however, become established that there are sex differences in a variety of abilities, such that males obtain higher mean scores on some abilities while females obtain higher mean scores on others. There is a large research literature on this issue including books by Kimura (1999), Halpern (2000) and Caplan, Crawford, Hyde, and Richardson (1997). Kimura (2002) lists five abilities on which males obtain higher means on average than females: spatial orientation, visualization, line orientation, mathematical reasoning, and throwing accuracy; and five abilities on which females obtain higher means on average than males: object location memory, perceptual speed, verbal memory, numerical calculation, and manual dexterity. To these should be added higher means obtained by females on spelling (Lynn, 1992) and foreign language ability (Lynn & Wilson, 1993), and higher means obtained by males on mechanical aptitude (Lynn, 1992).

The Wechsler tests provide one of the most useful sources of data for the study of sex differences in cognitive abilities because they are highly respected, being the most used intelligence test in the United States (Watkins, Campbell, Nieberding, & Hallmark, 1995) and extensively used in many countries, are based on well drawn standardisation samples, and provide measures of a good range of abilities including several verbal abilities (Information, Arithmetic, Vocabulary, etc), spatial and visualization abilities (picture completion, block design and object assembly), immediate or working memory (digit span), and perceptual speed (coding). Scores on these tests are aggregated to give a Full Scale IQ, which is widely regarded as a measure of general intelligence and Spearman's *g* (e.g. Jensen, 1998). There are three versions of the tests for adults, children aged 6–16, and pre-school children aged 4–6, and four editions designated the Wechsler–Bellvue, the Wechsler, the Wechsler-Revised, and the Wechsler-111. These tests have been standardised in many countries and therefore provide a rich source of data for the examination of sex differences in different abilities. It is asserted by Halpern (2000, p. 91) that "the overall IQ score does not show sex differences". As the overall score or full scale IQ is widely regarded as a good measure of general intelligence, this assertion is consistent with the position that there is no sex difference in general intelligence. However, Halpern's assertion has been disputed by Lynn (1994) who lists several studies in which males obtain higher Full Scale IQs than females by between 1 and 5 IQ points. This raises the problem of whether the Wechsler Full Scale IQ is a satisfactory measure of general intelligence or whether it is biased in favour of males. If it is a satisfactory measure of general intelligence, the existing evidence summarized by Lynn (1994) suggests that males have slightly higher general intelligence than females, contrary to the widespread consensus that that there is no sex difference in general intelligence.

To provide further data on sex differences in the abilities measured by the Wechsler tests and on the Full Scale IQ we present the results of a large sample of New Zealand children assessed on the WISC-R (Wechsler, 1974) at age 8 and 9 years.

2. Method

The data were gathered during the course of the Christchurch Health and Development Study (CHDS). The CHDS is a longitudinal study of a birth cohort of 1265 children born in the Christchurch urban region during mid 1977. This cohort has been now studied throughout childhood and into adulthood.

At ages 8 and 9 years, members of the cohort who were resident in the Canterbury region were assessed on the Revised Wechsler Intelligence Scale for Children (Wechsler, 1974). At age 8 years 881 children were studied and at age 9 years 811 children were studied. These samples represented between 64% and 70% of the original cohort of 1265 children. The reasons for restricting assessments to the Canterbury region were to reduce the overall costs of testing. However, comparisons between those included in the testing and those excluded showed that any bias due to socio-economic, educational or related biases was likely to be inconsequential (Fergusson, Fergusson, Horwood, & Kinzett, 1988; Fergusson & Lloyd, 1991).

The WISC-R was administered individually at school by trained psychometric testing staff. Testing was conducted strictly following the guidelines in the test manual but was limited to 8 of the 10 WISC-R sub-tests: Information; Similarities; Arithmetic; Vocabulary; Picture Completion; Block Design; Object Assembly and Coding. Each sub-test was scored using the criteria stated in the test manual and overall Verbal IQ, Performance IQ and Total IQ scores were obtained by pro-rating the sub-test scores using the formulae described in the test manual. Test reliability was assessed using split half methods and this showed the Verbal, Performance and Total IQ scores to have split half reliabilities that ranged between .87 and .95 (Fergusson et al., 1988).

For the purposes of the present analysis the sub-test and IQ scores at ages 8 and 9 were averaged across the two measurement periods. This approach has the advantage of reducing the amount of tabular material needed to display the results. However, the results were also confirmed by the use of a repeated measures analysis of variance.

3. Results

Table 1 shows mean scores on the sub-tests of the WISC-R for boys and girls. The Table also shows mean Verbal, Performance and Full Scale IQs. As explained in Section 2 all scores were averaged across 8 and 9 years and the analysis is based on all children who were observed at either 8 or 9 ($N = 897$). Each comparison is tested for statistical significance using the t test for independent samples. The size of the difference between groups is described by the d score. The d score is the difference between the two means divided by the pooled estimate of the standard deviation. Positive ds denote higher means of boys and negative ds denote higher means of girls.

The Table shows that there were significant ($p < .05$) gender differences in 5 of the 8 WISC-R sub-tests. In four of these comparisons (Information, Vocabulary, Block Design, Object Assembly) the differences favoured males with d scores ranging from .13 to .23. However, the male advantage in these tests was offset by a large female advantage ($d = -.53$) on the Coding sub-test. Females also showed small but non-significant advantages on the verbal sub-tests of Similarities and Arithmetic. To test the overall significance of sub-test differences, an overall multivariate test using Hotelling's T was conducted on the joint distribution of subtest scores between males and females. This showed a highly significant gender difference in sub-test means ($T^2 = 22.5$; df = 8887; $p < .0001$).

Table 1
Comparison of male and female test scores averaged over 8 and 9 years (minus signs in the *d* column denote higher means of girls)

Test	Mean (SD)		p[a]	d
	Boys (N = 448)	Girls (N = 449)		
Verbal				
Information	10.7 (3.2)	10.2 (2.9)	<.01	.18
Similarities	10.5 (3.6)	10.7 (3.4)	>.35	−.06
Arithmetic	9.3 (2.9)	9.5 (2.8)	>.50	−.04
Vocabulary	10.5 (3.1)	10.1 (2.9)	<.05	.13
Performance				
Picture completion	10.7 (2.4)	10.5 (2.3)	>.25	.07
Block design	11.6 (3.3)	10.9 (2.3)	<.01	.19
Object assembly	11.8 (2.7)	11.1 (2.7)	<.001	.23
Coding	8.7 (2.5)	10.1 (2.6)	<.001	−.53
Verbal IQ	101.5 (17.3)	100.5 (16.3)	>.35	.06
Performance IQ	104.6 (14.6)	104.6 (14.4)	>.90	.00
Full Scale IQ	103.1 (15.9)	102.5 (15.5)	>.50	.04

[a] Based on t test for independent samples.

The net effect of the gender related advantages and disadvantages on specific sub-tests was that there were only small and non-significant gender differences on Verbal, Performance and Full Scale IQ scores.

The findings in Table 1 are placed in an international context by the comparisons shown in Table 2. This table gives *d* scores for each sub-test and overall IQ scores for studies conducted

Table 2
Sex differences on the WISC-R in New Zealand, Scotland, Netherlands and the USA expressed as *d* scores (minus signs denote higher means of girls)

Tests	New Zealand N = 897	Scotland N = 1400	Netherlands N = 2027	USA N = 2200
Information	.18**	.39***	.30***	.37***
Similarities	−.06	.08	.08*	.07
Arithmetic	−.04	.12*	.09*	.06
Vocabulary	.13*	.28***	.14***	.14**
Picture completion	.07	.19***	.14***	.15**
Block design	.19**	.16**	.12**	.15**
Object assembly	.23***	.21***	.17***	.18**
Coding	−.53***	−.55***	−.36***	−.53***
Verbal IQ	.06	.22***	.11*	.19**
Performance IQ	.00	.01	.05	.01
Full scale IQ	.04	.13*	.09*	.12*

* $p < .05$.
** $p < .01$.
*** $p < .001$.

with the WISC-R in New Zealand, Scotland, the Netherlands and the USA. Overall, the table shows considerable agreement between the studies in that they all find significant gender differences for 5 of the eight WISC-R sub-tests with four of these differences (Information; Vocabulary; Block Design; Object Assembly) favouring males and one (Coding) favouring females. However, there are also some differences between the New Zealand results and findings from other studies. First, significant gender differences in Picture Completion were not found for New Zealand, whereas these were found in other studies. Second, gender differences in Information were less marked in the New Zealand sample. The net effects of these differences were that differences in overall IQ scores were smaller in the New Zealand sample than in the other samples. Whilst the New Zealand results show no significant gender differences in overall IQ scores, all other studies report a small male superiority in both Verbal and Full Scale IQ scores. The intercorrelations between the *d*s in the four countries are shown in Table 3.

To examine whether there is a sex difference in *g* we have adopted the method proposed by Jensen (1998, p. 538). This involves: (a) estimating an orthogonal *g* factor using Schmid–Leiman factor analysis, and (b) correlating sex with the *g* factor. This showed a small non-significant association ($r = .06$; $p = .09$) between sex and *g*. However, the Schmid–Leiman model showed a very poor fit to our data (LR $\chi^2(15) = 167.3$, $p < .0001$; RMSEA = .11; SRMR = .049; CFI = .96), raising questions about the suitability of this method for estimating the association between sex and general intelligence.

We have therefore adopted another method for examining whether the *r* is a sex difference in *g*. The results given in Tables 1 and 2 suggest the presence of a general trend for WISC-R scores to vary between boys and girls on both sub-test and total IQ scores. These sex differences may be explained in a number of ways. First, it may be suggested that the sex differences in test scores could imply that the factor structure of the WISC-R varied with gender. If this were the case then comparisons of total IQ scores of boys and girls may not be valid. To examine this possibility, a multiple group model of the structure of the WISC-R was fitted to the variance–covariance matrix of the eight sub-tests for males and females (see Table 4). The multiple group model is depicted in Fig. 1. This figure shows models for males and females fitted to a hierarchical model structure in which: (a) the observed sub-test scores were indicators of the first order verbal and performance factors; (b) the verbal and performance factors were predicted from a higher order factor of general intelligence. This model was fitted to the data in Table 4 in two forms: (a) an unconstrained model in which all coefficients were permitted to vary across gender groups; (b) a constrained form in which all factor loadings and factor variance parameters were equated across gender groups. These models were fitted using LISREL 8 (Joreskog & Sorbom, 1993) and methods of maximum likelihood estimation. In each case, the goodness of fit of the fitted model was assessed

Table 3
Between country correlations of *d* scores on WISC-R sub-tests[a]

Country	New Zealand	Scotland	Netherlands	USA
New Zealand	1.00			
Scotland	.94	1.00		
Netherlands	.94	.99	1.00	
USA	.94	.99	.99	1.00

[a] All correlations $p < .0001$.

Table 4
Matrix of correlations of 8 WISC-R sub-tests for males (above diagonal) and females (below diagonal)

	Information	Similarities	Arithmetic	Vocabulary	Picture completion	Block design	Object assembly	Coding
Information	1.00	.71	.64	.72	.52	.40	.33	.40
Similarities	.74	1.00	.54	.74	.46	.48	.35	.25
Arithmetic	.67	.59	1.00	.61	.38	.48	.35	.46
Vocabulary	.77	.76	.61	1.00	.49	.50	.39	.33
Picture completion	.48	.47	.40	.48	1.00	.48	.45	.37
Block design	.51	.55	.59	.51	.48	1.00	.65	.40
Object assembly	.42	.45	.44	.43	.52	.65	1.00	.39
Coding	.33	.28	.47	.30	.25	.40	.31	1.00

on the basis of a range of indices including: the log likelihood chi square statistic; the root mean squared error of approximation (RMSEA); the standardised root mean squared residual correlation (SRMR); and the comparative fit index (CFI). Examination of both models produced evidence of adequate fit (unconstrained model: $LR\chi^2(16) = 18.6$, $p = .29$; RMSEA = .019; SRMR = .009; CFI = 1.00) (constrained model: $LR\chi^2(25) = 26.1$, $p = .40$; RMSEA = .010; SRMR = .024; CFI = 1.00). The difference in model chi square statistics between the constrained and unconstrained models provides a test of the equivalence of the factor structure for males and females. This test was non-significant ($LR\chi^2(9) = 7.5$, $p = .59$), suggesting that the factor models for males and females were not different. The models depicted in Fig. 1 show the within group standardised model parameters for males and females from the constrained model solution. (It should be noted that to produce well-fitting models it was necessary to permit the errors of some sub-tests to be correlated. These residual correlations were typically small to moderate in size (range .10–.25), and suggested evidence of correlated test specificity that was independent of the factor structure for verbal, performance and general intelligence. The disturbance correlations that accounted for the largest improvement in model fit were between the errors for: object assembly and block design; coding and arithmetic; and block design and arithmetic. To simplify model presentation these error correlations have been omitted from Fig. 1.)

An alternative explanation is that there may be gender differences in: (a) general intelligence; or (b) sub-test performance independently of general intelligence. To examine this possibility the structural equation model shown in Fig. 2 was fitted to the data. This model permits sex to have a direct influence on general intelligence via the path from sex to general intelligence, or to have specific effects on sub-test performance independently of general intelligence via the paths from sex to the sub-tests. It should be noted that to identify this model sex can directly influence only 7 of the 8 sub-tests. However model fitting showed that the best fitting and most parsimonious model involved paths from sex to five sub-tests (information, vocabulary, block design, object assembly and coding). The model was fitted to the correlation matrix of the 8 sub-tests for the total sample, again using LISREL 8 and methods of maximum likelihood estimation. Goodness of fit indices for the fitted model suggested an adequate fit to the observed data ($LR\chi^2(11) = 13.0$, $p = .29$; RMSEA = .014; SRMR = .011; CFI = 1.00). Again, in order to obtain a well fitting model it was necessary to permit a similar set of error correlations to those used for the models in Fig. 1. Inspection of Fig. 2 leads to the following conclusions:

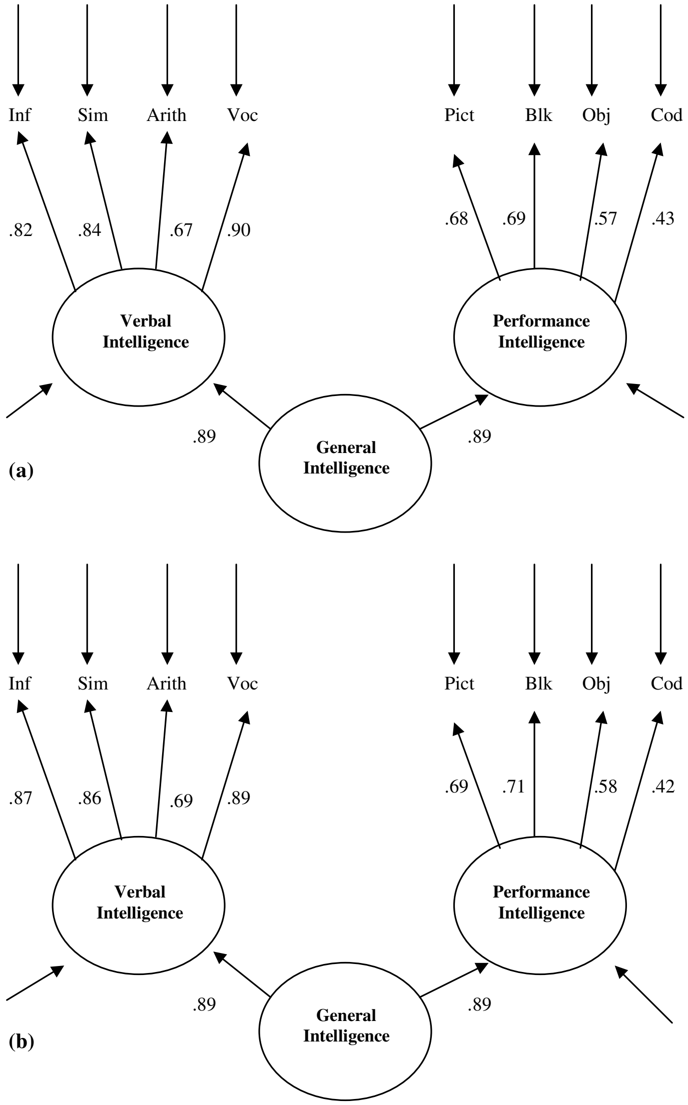

Fig. 1. Fitted hierarchical factor models for (a) males and (b) females.

1. Sex had no direct effect on general intelligence, suggesting that both males and females had the same mean levels of "*g*".
2. There was clear evidence of specific pathways from gender to test performance on 5 WISC-R sub-tests. For four of these tests (information, vocabulary, block design, object assembly) there was evidence that independently of general intelligence males performed better than females, whereas females showed clear superiority of performance on coding. These results, of course, mirror the differences shown in Table 1.

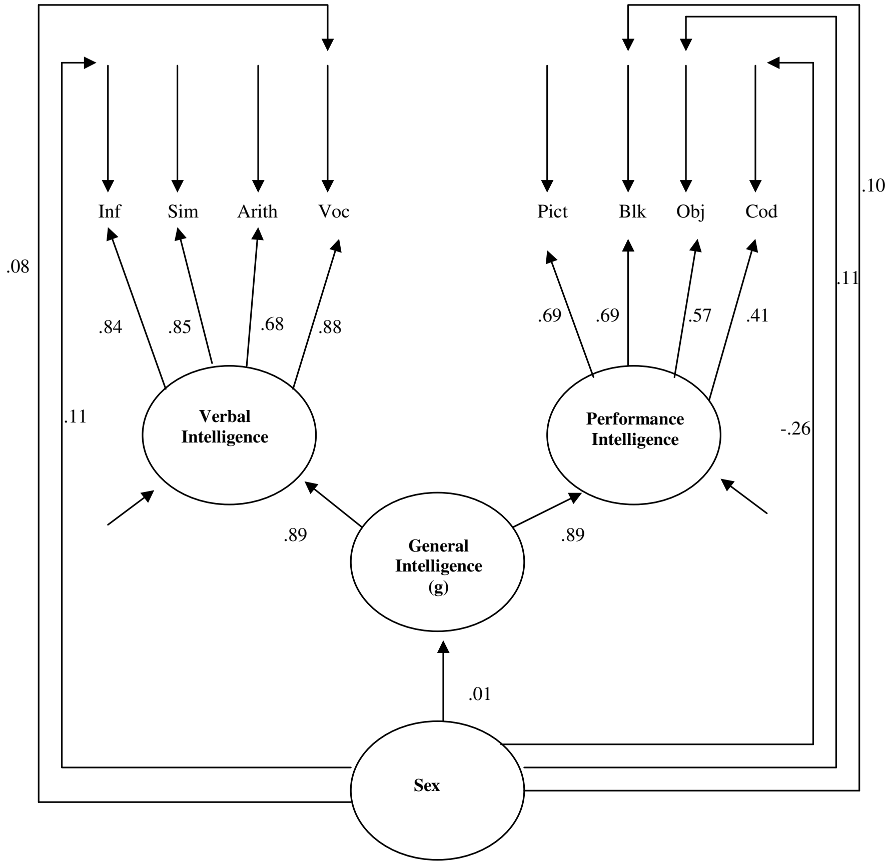

Fig. 2. Model of the effects of sex on general intelligence (g) and sub-test specificities.

The implication of the results in Fig. 2 is that although gender did not influence general intelligence, there were gender differences in specific abilities that led males to perform slightly better on tests of information, vocabulary, block design and object assembly, and females to perform substantially better on coding. These findings suggest that the gender differences in overall WISC-R scores did not reflect gender differences in general intelligence but rather reflected the presence of test invalidity in the overall WISC-R scores, with this invalidity arising from sex differences in the specificities of the WISC-R sub-tests.

4. Discussion

The sex differences in cognitive abilities on the WISC-R found in this New Zealand sample are closely similar to those in Scotland, the Netherlands and the United States. This is shown in the high correlations between the magnitude of the differences on the sub-tests expressed as *d*s given in Table 3. These range between .94 and .99 and are highly statistically significant. These very high correlations attest to the robustness and replicability of gender differences in narrow abilities across different cultures. In considering comparisons between the sex differences in the present sample and in the three other countries it should be borne in mind that the samples from Scotland, the Netherlands and the USA are for the age range 6–16 years, while the New Zealand sample consists of 8–9 year olds who are therefore younger than the mean age of 11 years of the other three samples.

Looking at the results in detail, there are eight principal points of interest in these New Zealand data. First, the boys obtained a significantly higher mean on Information (.18*d* = 2.7 IQ points), although the boys' advantage is not so large as in the other three countries (.39*d*, .30*d*, .37*d*). These results confirm those of several studies of the Wechsler information tests among adults and of other studies finding that among adults men have significantly higher means than women on information and general knowledge (Lynn & Irwing, 2002; Lynn, Irwing, & Cammock, 2002). It is interesting to find that the boys' advantage on information is present among 8 and 9 year olds, since in general sex differences among adults are smaller and sometimes non-existent among pre-adolescent children (Lynn, 1994). None of the recent books on sex differences in cognitive abilities by Caplan et al. (1997); Kimura (1999) or Halpern (2000), or the textbook on intelligence by Brody (1992) or Mackintosh (1998), makes any mention of a sex difference in general knowledge. We believe that this is now sufficiently well established to be added to the list of sex differences in cognitive abilities given in Section 1.

Second, the boys in the New Zealand sample have a significantly higher mean on Vocabulary. The size of the boys' advantage in New Zealand (.13*d*) is not so great as that in Scotland (.28*d*) but is almost exactly the same as that in the Netherlands (.14*d*) and the United States (.14*d*). This is contrary to the conclusion reached by Hyde and Linn (1988) (and endorsed by Caplan et al., 1997), in their meta-analysis of sex differences in verbal abilities, in which they calculated that there is a (negligible) female advantage of .02*d* on vocabulary. There is no apparent explanation for this discrepancy, but this meta-analysis cannot be regarded as providing a full coverage of studies because it did not include the standardisation samples of any of the Wechsler tests which are among the largest and best drawn samples.

Three, the New Zealand results show a negligible sex difference (−.06) in Similarities, while boys have a slight advantage in Scotland, the Netherlands and the United States, although the difference is only statistically significant in the Netherlands.

Four, in Arithmetic also the New Zealand results show a negligible sex difference (−.04), while boys have a slight advantage in the other three countries which is statistically significant in Scotland and the Netherlands, but not in the United States.

Five, for the verbal tests as a whole, New Zealand boys have a small but not statistically significant higher verbal IQ by 0.06*d*. The small advantage of the New Zealand boys is less than in the Netherlands (.11*d*), Scotland (.22*d*) and the USA (.19*d*) (the verbal IQs shown in Table 2 for Scotland, the Netherlands and the United States have been calculated for the complete test and include the comprehension and digit span subtests). The trend of the results in the four countries is for boys to have slightly higher verbal ability. This is contrary to the conclusions reached by Hyde and Linn (1988) in their meta-analysis of sex differences in verbal ability, in which they calculated that in studies published before 1973 females obtained a higher mean by .23*d* (3.5 IQ points) and in studies published after 1973 females obtained a higher mean by .10*d* (1.5 IQ points). Hyde and Linn secured their results by including data on spelling and word fluency as verbal abilities, on which females obtain higher means (Lynn, 1992), and by excluding the verbal test of the Scholastic Aptitude Test (SAT), as well as the Wechsler standardisation samples.

Six, turning now to the performance subtests, the New Zealand boys have a non-significant advantage on Picture Completion (.07), while their higher means in Scotland, the Netherlands and the United States are all statistically significant. On Block Design and Object Assembly boys obtain significantly higher means (.19 and .23, respectively) of about the same magnitude in all four counties. These are both tests of visualisation or spatial ability, and these results are consistent with many studies finding that on this ability males obtain higher means, as shown in the meta-analyses carried out by Linn and Peterson (1985) and Voyer, Voyer, and Bryden (1995). In the New Zealand data, the male advantage of the two tests together is .21*d*, and is about the same as that in Scotland, the Netherlands and the United States.

Seven, the New Zealand results confirm those in Scotland, the Netherlands and the United States in showing that girls obtain a substantially and significantly higher mean on the Coding subtest. The girls' advantage of .53*d* is the largest sex difference in all four countries and the size of the girls' advantage in New Zealand is closely similar to that of .55*d* in Scotland and .53*d* in the USA; the girls' advantage in the Netherlands is only a little smaller at .31*d*. Coding is generally considered to be a measure of perceptual speed on which females typically out-perform males (Lynn, 1992; Kimura, 1999). In the Wechsler tests, Coding is aggregated with the other non-verbal tests to give the Performance IQ, and because girls perform much better than boys on Coding while boys perform a little better on the other non-verbal tests, the net result is that there is very little sex difference on the Performance IQ in any of the four countries. In recent years it has become increasingly recognised that combining these disparate tests to give a Performance IQ is not a useful procedure because Coding measures a different ability from the remaining tests, which are largely measures of visualisation and spatial abilities.

Eight, on the Full Scale IQ the New Zealand boys obtained a slightly but not significantly higher mean by .04*d* (0.6 IQ points). The sex difference expressed as *g* is also negligible in this sample. This confirms the conclusion reached by Jensen (1998, p. 539) that the Full Scale IQ of the Wechsler tests is a good approximate measure of *g*. Nyborg (2003) has also reported a small but not significantly higher mean on *g* in children in a sample in Denmark. The findings from the series of studies reviewed in this paper all suggest that males scored higher on the WISC-R than females, with these differences being significant for data from Scotland, the US and the Netherlands. As noted in Section 3, these sex differences in WISC-R scores may arise in a number of ways. First it is possible that there could be sex differences in the factor structure of the WISC-R calling the validity of the comparisons into question. Analysis of the New Zealand data did not support this conclusion and suggested that the factor structure of the WISCR was the same for males and females. Second, it could be suggested that these differences reflect the fact that males have slightly higher general intelligence than females. This conclusion was not supported by either the Schmidt Lieman analysis proposed by Jensen (1998) or by the structural equation model in Fig. 2. Both analyses suggested that sex was not significantly correlated with *g*. Finally it may be proposed that the higher WISC-R scores for males reflect sex related sources of test specificity in which males tend to score higher on some WISC-R subtests independently of general intelligence. The model in Fig. 2 is consistent with this explanation and suggests that, independently of general intelligence, males performed better on Information, Vocabulary, Block Design and Object Assembly, while females performed better on Coding. The implication of these findings is that the WISC-R total score contains a small bias which makes it appear that males have slightly greater general intelligence than females, with this bias arising from sex differences in performance on a number of WISC-R subtests.

Acknowledgement

This research was funded by grants from the Health Research Council of New Zealand, the National Child Health Research Foundation, the Canterbury Medical Research Foundation and the New Zealand Lottery Grants Board.

References

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

Caplan, P. J., Crawford, M., Hyde, J. S., & Richardson, J. T. E. (1997). *Gender differences in human cognition*. Oxford, UK: Oxford University Press.

Fergusson, D. M., Fergusson, J. E., Horwood, L. J., & Kinzett, N. G. (1988). A longitudinal study of dentine lead levels, intelligence, school performance and behaviour: Part II: Dentine lead and cognitive ability. *Journal of Child Psychology and Psychiatry, 29*, 793–809.

Fergusson, D. M., & Lloyd, M. (1991). Smoking during pregnancy and effects on child cognitive ability from the ages of 8 to 13 years. *Paediatric and Perinatal Epidemiology, 5*, 189–200.

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

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: a meta-analysis. *Psychological Bulletin, 104*, 139–155.

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

Joreskog, K. G., & Sorbom, D. (1993). *LISREL 8: structural equation modelling with the SIMPLIS command language*. Chicago, IL: Scientific Software International.

Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: MIT Press.

Kimura, D. (2002). Sex hormone influence human cognitive pattern. *Neuroendocrinology Letters, 23*(Suppl. 4) (Special Issue).

Linn, M. C., & Peterson, A. C. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. *Child Development, 56*, 1479–1498.

Lynn, R. (1992). Sex differences on the differential ability test in British and American adolescents. *Educational Psychology, 12*, 101–106.

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

Lynn, R., & Irwing, P. (2002). Sex differences in general knowledge, semantic memory and reasoning ability. *British Journal of Psychology, 93*, 545–556.

Lynn, R., Irwing, P., & Cammock, T. (2002). Sex differences in general knowledge. *Intelligence, 30*, 27–40.

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.

Mackintosh, N. J. (1998). *IQ and human intelligence*. Oxford: University Press.

Nyborg, H. (2003). Sex differences in *g*. In H. Nyborg (Ed.), *The scientific study of general intelligence*. Amsterdam: Elsevier.

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

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial ability: a meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250–270.

Watkins, C. E., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26*, 54–60.

Wechsler, D. (1974). *Manual for the Wechsler intelligence scale for children–revised*. New York: Psychological Corporation.