Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Personality and Individual Differences 39 (2005) 683–688
PERSONALITY AND INDIVIDUAL DIFFERENCES
www.elsevier.com/locate/paid

Sex differences in 3 year olds on the Boehm Test of Basic Concepts: Some data from Mauritius

Richard Lynn [a,*], Adrian Raine [b], Peter H. Venables [c], Sarnoff A. Mednick [d]

[a] *University of Ulster, Coleraine, Northern Ireland, UK*
[b] *Department of Psychology and Neuroscience Program, University of Southern California, USA*
[c] *Department of Psychology, York University, UK*
[d] *Social Science Research Institute, University of Southern California, USA*

Received 4 November 2004; accepted 7 February 2005
Available online 23 May 2005

Abstract

Sex differences on the Boehm Test of Basic Concepts are examined in a sample of approximately 1400 3 year olds in Mauritius. Girls obtained a significantly higher Full Scale IQ by .11*d*, the equivalent of 1.65 conventional IQ points, and a significantly higher performance IQ by 1.95 IQ points. The results confirm previous studies that have found that among preschool children girls have slightly but significantly higher IQs than boys. There was no sex difference in variance, contrary to the frequent assertion that variance is greater in males than in females.

Keywords: Sex differences; Boehm test; Intelligence; Variance; Mauritius

1. Introduction

For approximately a century it has been consistently asserted that there is no sex difference in general intelligence. In the United States this conclusion was advanced by Terman (1916,

* Corresponding author. Tel.: +1 275 392 092.
E-mail address: lynnr540@aol.com (R. Lynn).

0191-8869/$ - see front matter © 2005 Published by Elsevier Ltd.
doi:10.1016/j.paid.2005.02.021

pp. 69–70) on the 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". A few years later Spearman (1923) in Britain asserted that there is no sex difference in *g*. In the second half of the twentieth century numerous authorities reaffirmed this conclusion. For instance, Cattell (1971, p. 131) concluded that "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". Brody (1992, p. 323) concluded that "gender differences in general intelligence are small and virtually non-existent"; Jensen (1998, p. 531) that "no evidence was found for sex differences in the mean level of *g*"; Mackintosh (1998, p. 567) that "there is no sex difference in general intelligence worth speaking of"; Lubinski (2000, p. 416) that "most investigators concur on the conclusion that the sexes manifest comparable means on general intelligence"; and Halpern (2000, p. 218) that "sex differences have not been found in general intelligence".

Despite this consensus that there is no sex difference in intelligence, it has occasionally been contended that among infants between the ages of 1 and 5 year girls are on average more cognitively advanced than boys and may therefore be said to have higher intelligence. An early statement of this view was advanced by Doran (1907, p. 425): "it is generally conceded that girls develop more rapidly in infancy. Boys talk but little under 24 months. This will account for the superior vocabularies of girls during the first few years". Many studies have found that vocabulary is a good measure of intelligence and vocabulary is assessed in virtually all intelligence tests that have a verbal component, including the various Wechsler, Stanford Binet and Kaufman tests. Doran (1907) reported observations that girls have larger vocabularies than boys by approximately 50% up to the age of 30 months. For instance, he found that at the age of 24 months, girls had an average vocabulary of 573 words while boys had an average vocabulary of only 367 words. His numbers, however, were very low at only 13 girls and 11 boys. Nevertheless a similar advantage of girls in vocabulary among two year olds was reported by Nelson (1973) and by Huttenlocher, Haight, Bryk, Seltzer, and Lyons (1991), who found that girls had a 13 words vocabulary advantage at 16 months, a 51 words advantage at 20 months and a 115 words advantage at 24 months. Once again, however, the numbers were very low at only 10 girls and 12 boys. A more recent study by Lutchmaya, Baron-Cohen, and Raggatt (2002) has again confirmed that among infants girls have larger vocabularies than boys by around 30%. They found that at 18 months, the girls' advantage was .64*d* and at 24 months it was .60*d*.

There have also been occasional reports that among pre-school children girls have higher IQs than boys. For instance, in the United States in the 1932 standardisation sample of the Stanford-Binet, among 2–5 year olds girls had a higher average IQ by 3.0 IQ points (McNemar, 1942). In the 1972 standardisation sample of the McCarthy Scales of Children's Abilities, 2–5 year old girls had a higher average IQ by 2.6 IQ points (Kaufman & Kaufman, 1973). Similar results have been found in Sweden using the Griffiths (1970) Development Scale in a longitudinal study of infants ($n = 452$), among whom it was found that at the age of one year girls were more advanced verbally than boys by 3.6 Development Quotient (DQ) points, and at 4 years ($n = 412$) girls were more advanced verbally than boys by 2.5 DQ points (Nordberg, Rydelius, & Zetterstrom (1991)).

These spasmodically reported results that among infants and preschool children between the ages of 1 and 5 year girls are on average more cognitively advanced than boys have not however become widely accepted. There is no mention of them in general books on intelligence such as those by Cattell (1971), Brody (1992), Caplan, Crawford, Hyde, and Richardson (1997), Jensen (1998) and Mackintosh (1998). They are mentioned in the specialist texts on sex differences in intelligence by Halpern (2000) and Kimura (1999) but not systematically reviewed. The present paper makes a contribution to this issue by reporting sex differences on the Boehm Test among 3 year olds in Mauritius.

2. Method

The participants were obtained from two towns (Vacoas and Quatre Bornes) in Mauritius, chosen because they were representative of the ethnic mix of the island. All children born between 1969 and 1970 in the two towns were recruited and participated in the study, for which their families were given incentives in the form of food supplies. The ethnic mix of the sample was 69% Indian, 26% predominantly Creole (largely of mixed African and European descent), and 6% other (Chinese, English, French and others). This almost exactly matches the ethnic proportions of the population found in census returns. At the age of 3 years the participants numbering approximately 1400 were tested with the Boehm Test of Basic Concepts-Preschool Version (Boehm, 1986) (the numbers differ slightly for different subtests and are given for each subtest in Table 1). Further details of the sample are given by Raine, Venables, Reynolds, and Mednick (2002).

The Boehm Test of Basic Concepts resembles the Wechsler tests in so far as it consists of a number of verbal and non-verbal subtests the scores of which can be summed to give verbal, performance and full scale IQs. A study by Li (2004) in which the Boehm Test and the WPPSI were both administered to a sample of 122 4 year olds obtained a correlation of .553 between IQs in the two tests. The tests took approximately 30 min to administer. Piloting of the Boehm test on three-year-olds indicated that some changes in format were necessary for age 3 testing. In addition to pilot tests in the laboratory, visits were made to the homes of pilot test children to observe them in a

Table 1
Gender differences in 3 year olds on the Boehm-3 Preschool test

Test	*N*	Boys Mean (SD)	*N*	Girls Mean (SD)	*t*	*d*
Arithmetic	739	10.05 (3.01)	686	9.93 (2.99)	.76	.04
Information	750	9.88 (3.02)	691	10.12 (2.98)	1.51	−.08
Color	744	9.90 (3.02)	689	10.16 (2.97)	1.28	−.09
Copying shapes	755	9.83 (2.89)	702	10.15 (3.01)	2.02*	−.11
Block assembly	752	9.86 (2.93)	701	10.14 (3.07)	1.80	−.09
Classification	744	9.87 (2.99)	691	10.14 (3.00)	1.72	−.09
Verbal IQ	724	99.50 (15.19)	664	100.52 (14.62)	1.27	−.07
Perform IQ	752	99.03 (14.86)	701	100.94 (15.10)	2.43**	−.13
Full scale	722	99.19 (15.00)	663	100.82 (14.93)	2.02*	−.11

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

more natural context. This piloting led to minor modifications of the test for use with Mauritian children. For example, sugar cane sticks were used for judgments of length (Mauritius had a predominantly sugar-cane economy in 1972), local rocks were used for judgments of size, pictures of Mauritian children were used for identification of body-parts of children, and a tea-set was used to assess ability to follow directions (tea-drinking is ubiquitous in Mauritius).

The modified test had 6 components: (1) block assembly (making constructions from blocks, e.g. bridge, circle, tower); (2) copying shapes (copying circle, triangle, and square); (3) information (identifying body parts, pictures of boys and girls); (4) number/size/length concepts (simple numerical ability, size and length discriminations); (5) color concepts (naming and pointing to different colors); (6) classification (making discriminations between same/different objects).

Most of these abilities parallel cognitive skills found in the Wechsler Preschool and Primary Scale of Intelligence (WPPSI, Wechsler, 1967) (e.g. labeling and Information, similar/different discriminations and Similarities, copying shapes and Geometric Design, number/size/length concepts and Arithmetic). Consequently, scale construction initially followed a face validity approach to form indices of verbal and spatial ability. Each scale was first normalized by transforming the raw scores to percentiles, and then finding the standard score for each percentile (Allen & Yen, 1979). Scales were then standardized to have a mean of 10 and standard deviation of 3.

Two of the scales (block assembly and copying shapes) were spatial-constructional in nature. The scores on these tests were summed and further standardized to a mean of 100 and SD of 15 to form an index of age 3 spatial ability. Coefficient alpha for this spatial scale was 0.46.

The remaining scales were verbal in nature. Some involved a verbal response (e.g. picture content, numbers) while others required verbal comprehension and knowledge of the names of objects (information). The scores on these tests were summed and standardized in the same way as the spatial tests to form an index of verbal ability. Coefficient alpha for the verbal scale was .76. Verbal ability correlated significantly with spatial ability ($r = .41$, $N = 1387$, $p < .0001$).

Verbal ability correlated significantly ($r = .31$, $p < .0001$) with a rating of the amount of verbalizations the child made to the experimenter (Raine, Reynolds, Venables, Mednick, & Farrington, 1998), while spatial ability correlated significantly with a measure of motor ability (jumping, hopping, balancing on one foot) at age 3 ($r = .23$, $p < .0001$). Data from 73 subjects who were given the Reynell Developmental Language Scale (Reynell & Huntley, 1972) at age 6 years showed a .36 correlation ($p < .002$) with the age 3 verbal ability measure compared to a .25 correlation ($p < .005$) with the spatial age 3 measure ($p < .025$); while the difference between these two correlations is not statistically significant due to the modest sample size, the verbal age 3 measure predicts twice the amount of variance in age 6 language than the spatial age 3 measure.

Confirmatory factor analysis was employed to assess whether the two-factor (verbal-spatial) model was a significantly better fit to the one factor model. A one-factor solution resulted in a significant mis-fit ($\chi^2(20) = 817.94$, $p < .0001$, RMSEA = .17). However, the two-factor model (Spatial, Verbal) resulted in a significant improvement in fit compared to the one-factor model ($\Pi^2(1) = 34.79$, $p < .0001$). The children were followed up at the age of 11 when 969 were tested with the WISC-R. The correlation between the IQ on the Boehm at age 3 and the full scale IQ at age 11 was .30 ($p < .001$) providing some validity for the Boehm as a test of general intelligence. The correlation is quite modest, consistent with other studies reviewed by Gottfried, Gottfried, Bathurst, and Guerin (1994) showing correlations of similar magnitude between IQs measured in preschool children and those obtained later.

3. Results

Sex differences on the six subtests, the Verbal IQ, the Performance IQ and the Full Scale IQ are shown in Table 1. This gives the means and standard deviations of the boys and the girls for each of the tests and for the IQs. This is followed by the *t* values for the statistical significance of the differences. In calculating the values of *t*, Levene's test for equality of variances was run and showed that the variances were not significantly different for any of the subtests. The *t*-tests were therefore computed assuming equality of variances. The right hand column gives the differences between the mean scores of the boys and girls expressed as *d*s (the differences between the means divided by the pooled standard deviations). Positive *d*s denote higher means obtained by boys and negative *d*s denote higher means obtained by girls. It will be seen that all the *ds* are negative except for that for arithmetic, showing that in general girls obtained higher mean scores.

4. Discussion

The results contain three points of interest. First, the girls obtained significantly higher means than the boys on all the subtests except arithmetic. The girls' advantage is quite small but is statistically significant at the $p < .05$ level for the copying shapes subtest and for the full scale IQ and at the $p < .01$ level for the performance IQ. The girls' advantage of .11*d* on the full scale IQ is equivalent to 1.65 conventional IQ points. The girls' advantage is therefore smaller than in the three previous studies of preschool children cited in the introduction, namely the girls' advantage of 3.0 IQ points among 2–5 year olds in the standardisation sample of the Stanford-Binet, 2.6 IQ points advantage of girls among 2–5 year olds in the standardisation sample of the McCarthy Scales of Children's Abilities, and the 2.5 DQ points in verbal ability among 4 year olds found in Sweden on the Griffiths Development Scale. Nevertheless, the present results confirm those of the previous studies in showing that among preschool children girls are more cognitively advanced than boys.

Second, a possibly surprising result is that the girls' advantage is greater on the performance IQ (.13*d*) than on the verbal IQ (.07*d*). This result is unexpected because the performance subtests are measures of visual-spatial abilities on which males at older ages have almost invariably been found to obtain higher average scores than females, as shown in the meta-analyses carried out by Linn and Peterson (1985) and Voyer, Voyer, and Bryden (1995).

Third, it may be interesting to note that the variances on the subtests are virtually identical for boys and for girls. On the first three subtests shown in the table, the variances are marginally greater for boys, while for the remaining three subtests, the variances are marginally greater for girls. The result that the variances are virtually identical for boys and for girls fails to confirm the claim frequently made that the variance of cognitive abilities is greater in males than in females. For instance: "the consistent story has been that men and women have nearly identical IQs but that men have a broader distribution...the larger variation among men means that there are more men than women at either extreme of the IQ distribution" (Herrnstein & Murray, 1994, p. 275); "males are more variable than females" (Lehrke, 1997, p. 140); "males' scores are more variable on most tests than are those of females" (Jensen, 1998, p. 537). Evidently this is not the case among 3 year olds in Mauritius.

References

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.

Boehm, A. (1986). *Boehm test of basic concepts—Preschool version*. San Antonia, Texas: Psychological Corporation.

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.

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

Doran, E. W. (1907). A study of vocabularies. *Pedagogical Seminary, 14*, 401–438.

Gottfried, A. W., Gottfried, A. E., Bathurst, K., & Guerin, D. W. (1994). *Gifted IQ: early developmental aspects*. New York: Plenum.

Griffiths, R. (1970). *The abilities of young children*. London: Child Development Research Center.

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

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

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: relation to language input and gender. *Developmental Psychology, 27*, 236–248.

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

Kaufman, A. S., & Kaufman, N. L. (1973). Sex differences on the McCarthy scales of children's abilities. *Journal of Clinical Psychology, 29*, 362–365.

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

Lehrke, R. (1997). *Sex linkage of intelligence*. Westport, CT: Praeger.

Li, X. (2004). Personal communication.

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

Lubinski, D. (2000). Scientific and social significance of assessing individual differences. *Annual Review of Psychology, 51*, 405–444.

Lutchmaya, S., Baron-Cohen, S., & Raggatt, P. (2002). Foetal testosterone and vocabulary size in 18- and 24-month old infants. *Infant Behavior and Development, 24*, 418–424.

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

McNemar, Q. (1942). *The revision of the Stanford-Binet scale*. Boston: Houghton Mifflin.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38*(149).

Nordberg, L., Rydelius, P., & Zetterstrom, R. (1991). Psychomotor and mental development from birth to four years: sex differences and their relation to home environment. *Acta Paediatrica Scandinavia, 378*(Suppl.), 1–25.

Raine, A., Reynolds, C., Venables, P. H., Mednick, S. A., & Farrington, D. P. (1998). Fearlessness, stimulation-seeking, and large body size at age 3 years as early predispositions to childhood aggression at age 11 years. *Archives of General Psychiatry, 55*, 745–751.

Raine, A., Venables, P. H., Reynolds, C., & Mednick, S. A. (2002). Stimulation seeking and intelligence: a prospective longitudinal study. *Journal of Personality and Social Psychology, 82*, 663–674.

Reynell, J., & Huntley, R. M. (1972). New scales for the assessment of language development in young children. *Journal of Learning Disabilities, 4*, 549–557.

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

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.

Wechsler, D. (1967). *Wechsler preschool and primary scale of intelligence*. San Antonio: The Psychological Corporation.