

SEX DIFFERENCES IN INTELLIGENCE: SOME COMMENTS ON MACKINTOSH AND FLYNN

RICHARD LYNN

University of Ulster, Coleraine, Northern Ireland

Students of the problem of sex differences in intelligence will be indebted to Mackintosh and Flynn for providing useful further clarification of the issue of sex differences in intelligence. I have little to say about Flynn's paper except to note that it confirms my thesis that among adults males have a small advantage over females in abstract reasoning ability. Mackintosh and I have some common ground but there remain substantial disagreements.

Mackintosh and I agree that to answer the question we have to agree on a definition of general intelligence and on the appropriate tests to measure this agreed definition. We agree that we cannot define general intelligence as the score on intelligence tests.

We agree that many people have regarded the Wechsler tests as good measures of general intelligence and that males obtain consistently higher full-scale IQs on these tests. Nevertheless, we agree that this does not provide a definitive answer to the question of whether there are sex differences in general intelligence because the Wechslers are an arbitrary collection of tests and a different collection might be assembled to produce no sex difference or female superiority. Mackintosh believes it would be easy to do this whereas I believe it would be impossible. Mackintosh criticizes the tests on the grounds that they do not contain a measure of abstract reasoning, but if they did the male advantage would be increased. He also criticizes them on the grounds that they are weighted towards being measures of verbal ability but if this were redressed by adding tests of non-verbal ability the male advantage on the test would again be increased. The problem for sex equalitarians is to propose several tests which might be added to a new 'gender fair' Wechsler, on which females out-perform males and which would correct the existing bias. Mackintosh does not rise to this challenge.

A point about the Wechslers which has not been raised is that they were constructed to minimize sex differences (Jensen, 1998, p. 533). As males consistently out-perform females on the test, the test constructors must have eliminated a number of the items on which the male advantage was particularly large. Possibly this is why the tests are largely measures of the second-order verbal factor which, of the three major second-order factors (the other two being reasoning and spatial abilities), is the one on which the male advantage is least pronounced. Yet for all the test constructors' efforts to remove the male advantage on the tests, they failed to eliminate it. The reason for this is that they set themselves an impossible task. Because males out-perform females on any credible assembly of cognitive abilities, there is no way of constructing a credible intelligence test on which males do not obtain a higher mean IQ than females.

This was found by Kaufman when he constructed his adult intelligence test and could not avoid a male advantage on fluid intelligence of 2.3 IQ points (Kaufman & Horn, 1996). My conclusion on the consistent male superiority on the Wechslers is that while it does not provide a conclusive solution to the problem of sex differences in intelligence, it undoubtedly presents sex equalitarians like Mackintosh with a problem.

Mackintosh says that the issue of sex differences in general intelligence cannot be solved by taking a test, factor analysing it, identifying the first principal component as *g*, and calculating sex differences in *g*. This is Jensen's (1998) approach, but Mackintosh says that it is no good because different test batteries yield different *g*s and different sex differences in *g*. I agree. This is why I did not go down this road in my 1994 paper.

I believe our major point of disagreement lies in the definition of general intelligence. If we could agree on this there is rather little scope for disagreement on the data to be used for the assessment of sex differences in general intelligence. Thus, the major difference between us comes down to my proposal that the definition of general intelligence which commands most consensus is Gustafsson's model, in which general intelligence is the product of the three second-order group factors of *Gc* (verbal comprehension), *Gf* (reasoning) and *Gv* (visualization or spatial ability). I fitted American data to the model, which among adults gave a male advantage of 1.7, 2.1 and 7.5 IQ points respectively, averaged to 3.8. I then fitted data for five other countries which yielded similar results.

It is possible to criticize my model and data fitting on the grounds that the model is wrong, that the data fitted to the model were chosen inappropriately, or that the contribution of the three second-order factors to general intelligence should have been weighted differently. Mackintosh criticizes my methodology on all three grounds. To start with the last, I adopted the simplest procedure of weighting the three second-order factors equally. Mackintosh says that this 'betrays a serious misunderstanding of the nature of hierarchical models', that the three factors have different loadings on the general factor and that I should have weighted the sex differences on the three factors by the factor loadings. I do not accept this criticism. Just because I adopt Gustafsson's model of general intelligence does not oblige me to accept the particular loadings of the three second-order factors on the general factor obtained by Gustafsson in one particular study. A number of datasets support the model but of course no two studies produce precisely the same loadings of the three second-order factors on the general factor. Hence it is not sensible to use the loadings obtained in one particular study. Nevertheless, even if we adopt Mackintosh's preferred procedure the effect on the result is very small. The factor loadings obtained by Gustafsson are 0.76 for *Gc*, 0.80 for *Gv* and 1.0 for *Gf*. If we give the sex differences on the three factors equal weight we have a male advantage on general intelligence of 3.80 IQ points. If we follow Mackintosh and weight the three factors with Gustafsson's factor loadings the sex difference is reduced to 3.13 IQ points. Mackintosh's suggestion that I do not understand hierarchical factor analysis is reduced to a quibble about 0.67 of an IQ point. Nevertheless, I contend that weighting the three second-order factors equally is the most sensible method for the estimation of the magnitude the sex difference in general intelligence.

Mackintosh criticizes some of the data I fitted to the model. He admits that males score substantially higher than females on the spatial abilities and brushes this aside with 'it would surely be more sensible to acknowledge that males outscore females on

tests of spatial ability – and perhaps even attempt to understand why such a difference occurs'. The explanation of the difference is precisely the issue to which I devoted a section of my 1994 paper. I proposed that the higher male spatial abilities evolved because of male specialization in hunting and tool-making skills during the hunter-gatherer stage of human evolution. I have discussed the neurological basis of the higher male spatial abilities in Lynn (1987) in which I argue that more of the brain is devoted to the spatial abilities in males than in females. There is no need for Mackintosh to urge me to discuss the reasons for the male advantage in spatial abilities. I have already done so.

Mackintosh rejects the data I used for reasoning ability (discussed below) and my use of the verbal tests of the Wechsler standardization samples for the quantification of sex differences in the verbal comprehension group factor. He asks why I did not use the Hyde and Linn meta-analysis of sex differences in verbal abilities which, he says, produced a female advantage of 1.65 IQ points. There are four answers to his question. First, he misrepresents the results of this meta-analysis; the actual result of weighting all the studies by sample size was a male advantage of 0.6 IQ points (see the authors' Table 3). It is only by the arbitrary elimination of some of the results in which males scored higher than females that he manages to convert this into a female advantage of 1.65 IQ points. Secondly, the meta-analysis does a poor job of compiling all the relevant studies. It omits all five of the Wechsler studies I listed in my 1994 paper and the Project Talent data based on approximately 75,000 subjects, all of which show a male advantage (Wise, Steel & MacDonald, 1979). Thirdly, it omits data on numerical ability and short-term and long-term verbal memory, all of which are represented in the Wechsler verbal IQ. The basic problem with Mackintosh's position is that he takes a too narrow view of the verbal comprehension group factor. He thinks it consists only of vocabulary and reading ability. It is more comprehensive than this and the Wechsler verbal IQs provide a more representative measure of the factor. Fourthly, the Wechsler tests are based on meticulously drawn representative samples of the population whereas the data Mackintosh relies on are averages of a lot of poorly sampled studies. For these reasons the Wechsler data provide a better measure of the second-order verbal factor than the data favoured by Mackintosh.

Mackintosh declines to accept my use of the Gustafsson model of general intelligence. Nevertheless there is an emerging consensus that this is the best model. Thus Carroll (1993, p. 624) writes that 'There is abundant evidence for a factor of general intelligence found . . . in performing induction, reasoning, visualization, and language comprehension tasks'; and Jensen (1998, p. 89) 'The most *g*-loaded items are those that involve some form of inductive or deductive reasoning, problems that involve spatial visualization, quantitative reasoning, and verbal knowledge and reasoning'. No mention here of the central importance of minor abilities like spelling or perceptual speed which Mackintosh urges should be given equal weight with the three major second-order abilities in the estimation of general intelligence.

It is easy to see why Mackintosh will not accept the Gustafsson model of general intelligence. He realizes that the presence of spatial ability in the model is bound to give males a decisive advantage. Mackintosh's starting point is that there is no sex difference in general intelligence. Granted this premise, his main problem is that he has to find some credible way of massaging away the large male advantage on the spatial

abilities. He has two shots at solving this problem. The first is to swamp the male advantage on the spatial abilities by giving equal weight to several minor abilities like spelling, clerical speed and perceptual speed. He seems to realize that this has no credibility so he goes for another solution consisting of eliminating the spatial abilities entirely from the estimation of general intelligence. He proposes to estimate general intelligence solely from fluid intelligence defined as reasoning ability. Of course no-one will accept that general intelligence consists solely of reasoning ability. Everyone agrees broadly with Humphreys (1994, p. 180) when he writes that 'Intelligence is the acquired repertoire of all intellectual (cognitive) skills and knowledge available to a person at a particular point in time'. Everyone accepts this, with the unique exception of Mackintosh, although people disagree about the weighting to be given to different abilities in the estimation of general intelligence.

Mackintosh proposes that the best test of reasoning ability is Raven's Progressive Matrices. He says there is no sex difference on this test and concludes that he has proved his case. I have shown in my preceding paper that there are three adequately sampled studies of sex differences among adults on the Progressive Matrices and all of them show a male superiority which can be averaged to 5.2 IQ points. Mackintosh now says there are other studies in Court's list which do not show a male superiority among adults. I have read these studies and I can find none other than the three which come anywhere near to being acceptable in regard to their sampling. I invite Mackintosh to reveal which studies he believes provide persuasive evidence that adult males do not out-perform females.

Mackintosh tries to discredit the three studies showing male superiority by arguing that sex differences in intelligence show cohort effects such that the difference among older generations has been eroding among the younger. He asserts that the evidence for this is strong. This is far from the case. On the Wechsler standardization samples the advantage of males has increased over time. In the prototype of the test, the Wechsler-Bellevue standardized in the 1930s, females obtained a higher full-scale IQ than males. In all subsequent standardization samples males have obtained higher full-scale IQs than females. The most recent standardization of the adult Wechsler test has been in Scotland, where males out-perform females by approximately 5 IQ points, the greatest sex difference on the Wechslers yet reported (Lynn, 1998).

There are no cohort differences in the male superiority in the Japanese standardization sample of the WAIS-R spanning the age range 17-74 years (Hattori & Lynn, 1997) or in fluid intelligence in the standardization sample of the Kaufman Adult Intelligence Test spanning the age range 17-94 years (Kaufman & Horn, 1996). Mackintosh is chasing a chimera in his attempt to show a secular decline in the male advantage in general intelligence. For every study he cites showing a decline in the male advantage, I can cite another showing an increase or no change.

Mackintosh thought he had found the evidence to support his case that there is no sex difference in abstract reasoning ability in some Israeli military conscript data provided to him by James Flynn. In his lecture at the 1996 Galton conference, subsequently published in his 1996 paper, he stated that these unpublished data showed that there is 'no sex difference in performance on Raven's Matrices in Israeli 18 year olds'. Flynn has now published this evidence in the preceding paper and we find that Mackintosh has misrepresented it. We find the Israeli test was not the Progressive

Matrices but an adaptation of it. We find also that males out-perform females on the test, consistent with their out-performance in other well-sampled studies of abstract reasoning ability. There are problems in estimating the magnitude of the male advantage because large numbers of the female age cohort were not tested. Nevertheless, Flynn makes various assumptions to overcome this problem and offers three estimates of the sex difference consisting of a male advantage of 2.0, 1.7 and 1.4 IQ points. These are almost certainly underestimates because the test is very short with relatively low reliability, and this will reduce the real sex difference. Nevertheless, in my 1994 paper I estimated the male advantage on abstract reasoning at 2.4 IQ points. The Israeli data, weak though they are, confirm my estimate to within 1 IQ point. They certainly disconfirm Mackintosh's position that there is no sex difference in abstract reasoning ability. Even adopting Mackintosh's own definition of general intelligence, his preferred measure of it and his preferred dataset, the Israeli data disconfirm his contention that there is no sex difference and support my thesis that there is a sex difference favouring males.

Conclusion. Mackintosh rightly says that if we have to tackle the problem of sex differences we have to start with a definition of general intelligence. His most fundamental problem is that he defines general intelligence as consisting solely of reasoning ability. This definition will be universally rejected. General intelligence is much broader than this. Once this is accepted, and once it is accepted further that spatial abilities are a significant component of general intelligence, then adult males will inevitably obtain higher means than females because there is no credible way of finding other major abilities, or combinations of minor abilities, to compensate for the male advantage on spatial abilities.

References

- CARROLL, J. B. (1993) *Human Cognitive Abilities*. Cambridge University Press, Cambridge.
- HATTORI, J. & LYNN, R. (1997) Male-female differences on the Japanese WAIS-R. *Personal. individ. Diff.* **23**, 531-533.
- HUMPHREYS, L. G. (1994) Intelligence from the standpoint of a pragmatic behaviorist. *Psychological Inquiry* **5**, 179-192.
- JENSEN, A. R. (1998) *The g Factor*. Praeger, Westport, CT.
- KAUFMAN, A. S. & HORN, J. L. (1996) Age changes on fluid and crystallized ability for women and men on the Kaufman Adolescent and Adult Intelligence Test (KAIT) at ages 17-94 years. *Arch. clin. Neuropsychol.* **11**, 97-121.
- LYNN, R. (1987) The intelligence of the Mongoloids: a psychometric, evolutionary and neurological theory. *Personal. individ. Diff.* **8**, 813-844.
- LYNN, R. (1998) Sex differences on the standardisation sample of the Scottish WAIS-R. *Personal. individ. Diff.* **24**, 289-290.
- WISE, L. L., STEEL, L. & MACDONALD, C. (1979) *Origins and Career Consequences of Sex Differences in High School Mathematics Achievement*. American Institutes for Research, Palo Alto.