

INTELLIGENCE

Is there a sex difference in IQ scores?

Arising from: S. Blinkhorn *Nature* 438, 31–32 (2005)

Steve Blinkhorn¹ criticizes our study² of samples of university students, in which we found that the average IQ of men is 4.6 points higher than that of women, as measured by the Progressive (or Raven's) Matrices. He maintains that there is a negligible sex difference in adult intelligence. We believe that the principal error of Blinkhorn's criticism is that he does not consider our result in the context of several other studies showing that adult males have an IQ advantage of around 4–6 IQ points^{3–7}.

The strongest evidence for our position consists of a meta-analysis³ derived from 57 general-population samples, many of which are normative or carefully constructed representative samples, with a total of 80,928 participants. This reveals no sex difference in general cognitive ability up to the age of 14 and a significant sex difference at 15, which then increases to its adult value of 5 IQ points in favour of males. There is further evidence for a mean male advantage of 4–6 IQ points in four independent adult samples^{5–7} ($n = 11,896$); Jackson and Rushton⁸, in a huge standardization sample ($n = 102,515$) that is much bigger than the Mexico study⁹, also reported a male advantage of 3.6 IQ points among 17-year-olds, which is somewhat greater than our estimates^{3,4} for this age group.

At two points in his critique¹, Blinkhorn addresses the 'file drawer' problem, in which non-significant results fail to get published. Such a bias cannot be operating in this instance because, as Blinkhorn himself notes, almost none of the published literature on the Progressive Matrices has focused on sex differences. Furthermore, the idea that a null sex difference is unpublishable is belied by the huge number of books and articles, including Blinkhorn's, that claim exactly that. Therefore, the fact that 21 out of 22 student samples showed a male advantage suggests the phenomenon is robust.

Blinkhorn criticizes us for not adopting the principle of weighting results by sample size, and for excluding the very large study from Mexico⁹. This misses a central point of meta-analysis. We carried out a number of tests for moderator variables (factors that cause under- or overestimates of the sex difference) and found strong evidence for two: these were the type of test and the tendency of some universities selectively to recruit either brighter men or brighter women. In the presence of strong moderators, many of the studies in the sample provide biased estimates of the sex difference in IQ score.

It is clear from the box plot (Fig. 1) that the Mexico results conform to estimates from the

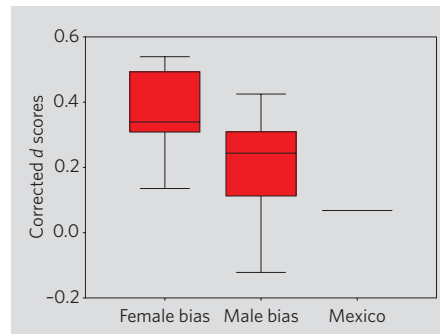


Figure 1 | Selection biases in IQ measurements. A comparison of Mexico, on the d -score (male mean minus the female mean, divided by the within-group standard deviation) difference in IQ, with samples showing pro-female and pro-male selection biases among university students. Horizontal bars represent the median; red rectangles, the interquartile range; tails, outliers. Sample sizes: female-biased studies, number of samples $k = 10$, $n = 6,812$; male-biased studies, $k = 10$, $n = 4,296$; Mexico study, $k = 1$, $n = 9,048$.

most male-biased samples, which provide substantial underestimates of the sex difference in IQ. Given the strong probability of bias in this sample, to weight it by its sample size (9,048) would risk a serious underestimate of the population sex difference in IQ. For this reason, we followed the advice of a definitive article on meta-analysis¹⁰ and took the median of estimates, including Mexico⁹, which equated to 4.6 IQ points.

Many of Blinkhorn's difficulties stem from his assumption that our focus was on university students. This makes little sense, because the IQ difference in students is dependent on which population is considered, whereas the sex difference in the general population, our actual focus of interest, is highly stable³. This also explains our choice of Cohen's d (Fig. 1),

as this provides a standard metric that partly controls for range restriction. According to accepted meta-analytical procedures¹¹, a difference of $0.31d$ translates into a 4.6-point IQ difference in the general population: this is an underestimate and not the overestimate suggested by Blinkhorn.

Blinkhorn also implies that the Progressive Matrices may be biased against women. This issue has been investigated¹² by using the conservative procedure of eliminating possibly biased items, identified by methods of differential item functioning: that analysis showed negligible evidence of bias.

However, we agree with Blinkhorn on the need for large representative samples, which is why we continue^{3,4} to make use of these.

Paul Irwing*, Richard Lynn†

*Manchester Business School, University of Manchester, Booth Street West, Manchester M15 6PB, UK

e-mail: paul.irwing@man.ac.uk

†School of Psychology, University of Ulster, Coleraine, Northern Ireland BT52 1SA, UK

1. Blinkhorn, S. *Nature* 438, 31–32 (2005).
2. Irwing, P. & Lynn, R. *Br. J. Psychol.* 96, 505–524 (2005).
3. Lynn, R. & Irwing, P. *Intelligence* 32, 481–498 (2004).
4. Lynn, R., Allick, J. & Irwing, P. *Intelligence* 32, 411–424 (2004).
5. Colom, R., Juan-Espinosa, M., Abad, F. J. & Garcia, L. F. *Intelligence* 28, 57–68 (2000).
6. Colom, R., Garcia, L. F., Juan-Espinosa, M. & Abad, F. J. *Span. J. Psychol.* 5, 29–35 (2002).
7. Nyborg, H. in *The Scientific Study of General Intelligence* (ed. Nyborg, H.) 187–227 (Elsevier, Amsterdam, 2003).
8. Jackson, D. N. & Rushton, J. P. *Intelligence* (in the press).
9. Backhoff-Escudero, E. *Rev. Mex. Psicol.* 13, 21–28 (1996).
10. Rosenthal, R. *Psych. Bull.* 118, 183–192 (1995).
11. Hunter, J. E. & Schmidt, F. L. *Methods of Meta-analysis: Correcting Error and Bias in Research Findings* 2nd edn (Sage, London, 2004).
12. Abad, F. J., Colom, R., Rebollo, I. & Escorial, S. *Pers. Ind. Diff.* 36, 1459–1470 (2004).

doi:10.1038/nature04966

INTELLIGENCE

Blinkhorn replies

Replying to: P Irwing & R. Lynn *Nature* 442, doi:10.1038/nature04966 (2006)

Irwing and Lynn explicitly claim to have identified a sex difference in IQ among university students¹. They now counter² my criticism³ of the flawed methodology upon which their conclusions are based and claim that an estimate of what this difference may actually be is irrelevant.

Irwing and Lynn's other work on the Progressive Matrices⁴ contains no actual statistics,

only estimates of "effect sizes", so it is not susceptible to the ordinary processes of scrutiny. The large study in press⁵ is not based on Raven's Matrices but on the American SAT college-entrance examination, has an unexplained excess of female participants, and raises several technical and substantive issues beyond the scope of my original critique.

Their response² to my critique, in which they

protest that they are complying with good meta-analytical practice, deflects attention from the inappropriate meta-analytical manoeuvres in which they have become entangled.

For the Standard Progressive Matrices, for instance, they report standard deviations from 3.02 to 9.69, and mean scores from 14.77 to 55.26, with no explanation for these extraordinary variations. They equate each of these widely differing standard deviations in turn to the conventional figure for IQ in the whole population, and then calculate the median of the resulting sex differences, each now in a different, exaggerated and artificial metric. Their meta-analytic paraphernalia are not needed, neither are they helpful, as the original scores themselves provided a suitable metric.

The Mexican study⁶ is not biased: rather, Irwing and Lynn's Fig. 1 shows how far removed they have become from the data. Their argument here is circular: the sex difference is vanishingly small compared with their sample of smaller, less representative groups, so therefore there must be a bias.

In fact, the Mexican undergraduates⁶ were

tested as part of the admissions process (which adds the crucial element of motivation that tends to eliminate sex differences) at a university that provides for 70% of the demand in its region, covering the whole range of courses. Four-and-a-half times as many people were tested as were tested in the other nine studies using the Standard Progressive Matrices combined.

There were marked IQ differences between gender-stereotyped disciplines (up to about two-thirds of a standard deviation, over four raw-score points), with engineering and physical science disciplines scoring highest, and social sciences and humanities lowest. But the difference between the sexes was only one-tenth as much, and described by the author as tiny ("minima")⁶.

Most large-scale ability testing goes unreported in academic journals. Those who engage with recruitment and selection campaigns rarely have the need, the motivation, or indeed (as in my case) the permission to report all they find. That is why there is a big file-drawer effect. The place to look for group differences is in the archives of institutional

test users, where one can also find the results of efforts to identify and eliminate possible ethnic and sex biases. That, and not a suggestion that Raven's Matrices are biased, was the import of my earlier remarks³.

I do not say that Irwing and Lynn's approach to meta-analysis is idiosyncratic, rather that their use of it here is methodologically inappropriate, statistically bizarre and unsuited to the estimation of subpopulation parameters. It has wrongly given rise to claims that IQ tests are biased against women and so should be abandoned.

Steve Blinkhorn

Psychometric Research & Development,
Brewmaster House, The Maltings,
St Albans AL1 3HT, UK
e-mail: steve@prd.co.uk

1. Irwing, P. & Lynn, R. *Br. J. Psychol.* **96**, 505-524 (2005).
2. Irwing, P. & Lynn, R. *Nature* **442**, doi:10.1038/nature04967 (2006).
3. Blinkhorn, S. *Nature* **438**, 31-32 (2005).
4. Lynn, R. & Irwing, P. *Intelligence* **32**, 481-498 (2004).
5. Jackson, D. N. & Rushton, J. P. *Intelligence* (in the press).
6. Backhoff-Escudero, E. *Rev. Mex. Psicol.* **13**, 21-28 (1996).

doi:10.1038/nature04967