J.biosoc.Sci, (2007) **39**, 789–793, © 2007 Cambridge University Press
doi:10.1017/S0021932007001861 First published online 2 Mar 2007

Short Report

PARTICIPATION RATES AND THE DIFFERENCE IN PERFORMANCE OF WOMEN AND MEN IN CHESS

MERIM BILALIĆ* AND PETER MCLEOD†

**School of Social Sciences and Law, Brunel University, Uxbridge, UK and*
†Department of Experimental Psychology, Oxford University, Oxford, UK

Summary. The superiority of men over women in chess has been cited as evidence that there are fundamental differences in male and female intelligence (Howard, 2005a, 2006; Irwing & Lynn, 2005). An alternative interpretation of the difference is that it is due to differential male and female participation rates in chess (Charness & Gerchak, 1996; Bilalić & McLeod, 2006; Chabris & Glickman, in press). This has been dismissed by Howard (2006) on the grounds that changes in the difference in skill level between top male and female players in recent years are not correlated with changing relative participation rates. Here it is shown that Howard's analysis is misleading. The data are consistent with differential participation rates as the explanation of the gap between the performance of women and men in chess.

One intellectual domain where the performance of men appears to be superior to that of women is chess. For example, there is only one woman among the top 200 chess players in the world (Bilalić, Smallbone, McLeod & Gobet, unpublished). While explanations such as socialization and different interests or gatekeeper effects have been proposed, some have seen it as evidence of differences in the intellectual abilities of men and women (e.g. Howard, 2005a, 2006; Irwing & Lynn, 2005). However, when looking for explanations of the difference between top women and men performers it is necessary to consider the different participation rates. If two samples come from the same population, the larger sample is likely to contain individuals representing more extreme values (Charness & Gerchak, 1996; Glickman & Chabris, unpublished). In competitive chess men outnumber women by about 17 to 1 (Bilalić, Smallbone, McLeod & Gobet, unpublished). It has been shown that this difference in participation rate alone is enough to explain the superiority of the best men players over the best women (Bilalić, Smallbone, McLeod & Gobet, unpublished). Howard (2006) has recently claimed that differential participation rates cannot explain the differences in

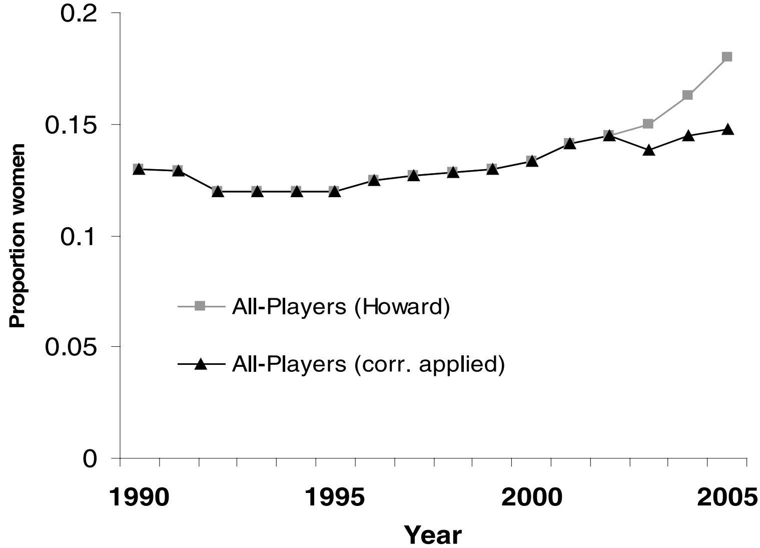

Fig. 1. The proportion of female players in the FIDE ranking lists from 1990 to 2005. Light squares are data taken from Howard (2005a), who included only men above Elo 2200. The cut-off for including female players was Elo 2000 from 1990 to 2003, Elo 1800 in 2003–4 and Elo 1600 in 2005. The black squares show the proportion when the criterion for including women remains constant at Elo 2000 from 2003 to 2005.

average ratings of women and men because changes in the difference are not associated with changes in relative participation rates.

Howard (2006) reported the proportion of women in the International Chess Federation (FIDE) ranking lists for players above a certain Elo rating (to be discussed below) for the years from 1975 to 2005 and the difference in the average performance of the male and female players during this period. (FIDE ranking lists include all players who achieve a minimum rating in tournaments recognized by FIDE. Elo rating is an interval scale representing the skill level of chess players with a mean of about 1500 and a standard deviation of about 200; Elo, 1978.) The data from 1975 to 1990 are noisy. Howard admitted (2005b, p. 350) that tracking down players in the pre-1990 lists had been a problem. Since 1990 FIDE has maintained its lists electronically and it is easier to be sure that the reported data are reliable. Figure 1 displays the relative proportion of women in the FIDE rankings from 1990 to 2005 reproduced from Howard's data. Figure 2 presents the difference between the average ratings of women and men in the cohort achieving the required Elo rating for inclusion (see figure legend).

From 1990 to 2003 there was a slow but steady increase in the proportion of women included in the rankings (Fig. 1) and a slow but steady decline in the average rating difference between men and women (Fig. 2). This is what would be predicted if the difference between top male and female players was caused by the much lower participation rate of women – when the proportion of women who achieve ratings over Elo 2000 goes up, the difference between the average level of the best male and

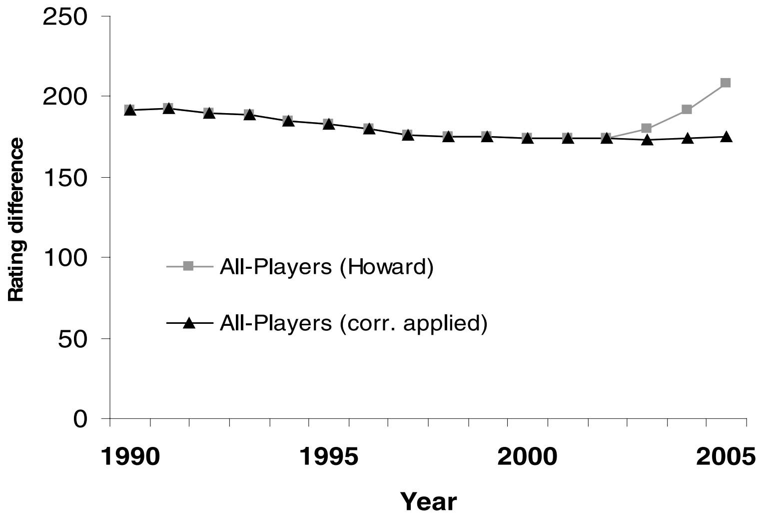

Fig. 2. Average difference in rating for FIDE rated men and women from 1990 to 2005 (see Fig. 1 for details of the criteria for inclusion.)

female players goes down. In 2003 there were abrupt increases in both the proportion of women playing and in the rating difference. According to Howard (2006) the post-2003 data demonstrate that participation rate does not determine rating difference because 'in the lowest participation years of 1993–1995, rating difference vary between 183 and 189 points' while 'in the highest, 2005, it is 208 points' (p. 423). If the rating difference was highest when the participation rate for women was highest, then participation rate cannot be the explanation for the difference between women and men's performance.

However, there is a simple explanation for the abrupt increase in rating difference and participation rates in the years 2003–5. In 2003, the minimum rating that female players needed to obtain to be listed by FIDE was lowered from 2000 rating points to 1800. In 2005 the minimum rating was lowered again to 1600. In the data reported by Howard (2005a) the weaker women included as a result of this rating change are included in the data but (for reasons described below) the corresponding weaker men are not! Hence, in Howard's data for 2003 and 4, women between 1800 and 2000 are included and then in 2005 women between 1600 and 1800 are included as well. This lowered the average rating for women in comparison with the previous years that included only women above 2000. But the men's average rating remained unaffected because Howard continued to use the cut-off of 2200 for inclusion. Howard (2005a) applied the correction of including only men over 2200 in order to compensate for the rating change that occurred in 1993 when the minimal rating for men to be included in the FIDE rating lists was dropped from 2200 to 2000. This produced a number of weaker male players who reduced the average, making it incomparable with the average ratings in previous years.

Thus in the data reproduced from Howard (2005a) in Fig. 2, between 2003 and 2005, a cohort of men of fixed strength are being compared with a cohort of women that includes more weaker players each year. Not surprisingly, the proportion of women increases during this period and the apparent superiority of male players increases! When the women included in the years 2003–5 are restricted to those who were eligible between the years 1990 and 2002 (as with the men), that is when the criterion for the inclusion of women is kept constant at Elo 2000, the rating difference did not increase between 2003 and 2005 (see the dark triangles in Fig. 2). And neither is there an increase in the number of women relative to men (see the dark triangles in Fig. 1).

Between 1990 and 2005, the years for which reliable data are available, the association between women's participation rates and rating differences between men and women (keeping male and female participation criteria fixed at Elo 2200 and 2000 respectively) is negative (around $r = -0{\cdot}60$ as estimated from Howard's (2006) figures). The more women that participate in chess at a level good enough to achieve an Elo score greater than 2000, the smaller the difference in rating between top men and women becomes. This result suggests that differences in chess skill between women and men may well be a result of differential participation rates.

Apart from the problems of changing criteria for inclusion in the male and female cohorts, another aspect of the data reported by Howard (2005a, 2006) prevents conclusions from being drawn about differences between average male and female players. The FIDE lists do not encompass 'data of a whole population' (Howard, 2005a, p. 371) but only of above-average players – and in the sub-sample of FIDE rankings reported by Howard, only very good male players. His cut-off values for men and women (ELO 2200 and 2000) were plus 3·5 SDs and 2·5 above the mean for men and women, respectively (theoretical mean and standard deviation of the Elo rating are around 1500 and 200, respectively; Elo, 1978). Conclusions drawn from an analysis of this elite sub-sample of players cannot be confidently extrapolated to the whole population of chess players. The FIDE ranking lists do not show whether the average male player is stronger than the average female player or not because the majority of players are not represented. In addition to this problem with FIDE ranking lists, failing to apply the correction for the minimal entry rating for women, the same correction applied for men in the previous paper, certainly does not add credibility to the claim that the gender differences in chess are partly a consequence of gender differences in intellectual ability.

References

Bilalić, M. & McLeod, P. (2006) How intellectual is chess? – A reply to Howard. *Journal of Biosocial Science* **38**, 419–421.

Chabris, C. F. & Glickman, M. E. (in press) Sex differences in intellectual performance: Analysis of a large cohort of competitive chess players. *Psychological Science*.

Charness, N. & Gerchak, Y. (1996) Participation rates and maximal performance: A log-linear explanation for group differences, such as Russian and male dominance in chess. *Psychological Science* **7**, 46–51.

Elo, A. E. (1978) *The Rating of Chess Players, Past and Present*. Arco, New York.

Howard, R. W. (2005a) Are gender differences in high achievement disappearing? A test in one intellectual domain. *Journal of Biosocial Science* **37**, 371–380.

Howard, R. W. (2005b) Objective evidence of rising population ability: A detailed examination of longitudinal chess data. *Personality and Individual Differences* **38**, 347–363.

Howard, R. W. (2006) IQ, visuospatial ability and the gender divide: a reply to Bilalić and McLeod. *Journal of Biosocial Science* **38**, 423–426.

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.