Why are (the best) women so good at chess?
Participation rates and
gender differences in intellectual domains

Merim Bilali, Kieran Smallbone, Peter McLeod and
Fernand Gobet

March 2009


Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from:  http://www.manchester.ac.uk/mims/eprints
And by contacting:  The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097
Why are (the best) women so good at chess? Participation rates and gender differences in intellectual domains

Merim Bilalić1,*, Kieran Smallbone2, Peter McLeod1 and Fernand Gobet3

1Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK
2Manchester Centre for Integrative Systems Biology, University of Manchester, 131 Princess Street, Manchester M1 7DN, UK
3Centre for Cognition and Neuroimaging, Brunel University, Uxbridge, Middlesex UB8 3PH, UK

A popular explanation for the small number of women at the top level of intellectually demanding activities from chess to science appeals to biological differences in the intellectual abilities of men and women. An alternative explanation is that the extreme values in a large sample are likely to be greater than those in a small one. Although the performance of the 100 best German male chess players is better than that of the 100 best German women, we show that 96 per cent of the observed difference would be expected given the much greater number of men who play chess. There is little left for biological or cultural explanations to account for. In science, where there are many more male than female participants, this statistical sampling explanation, rather than differences in intellectual ability, may also be the main reason why women are under-represented at the top end.

Keywords: gender differences; participation (base) rates; chess; intellectual activities; intelligence; science

1. INTRODUCTION

The former president of Harvard University, Lawrence Summers, expressed a view widely held by both academic researchers (e.g. Geary 1998; Kimura 1999; Pinker 2002) and laypeople when he suggested that innate biological differences in intellectual abilities may explain why the most successful scientists and engineers are predominantly male (Summers 2005). Recent research confirms that, although there is little difference between the average scores of men and women on intelligence and aptitude tests, highly intelligent people are predominantly male (Deary et al. 2003; Irwing & Lynn 2005). As Irwing & Lynn (2005) said: ‘different proportions of men and women with high IQs… may go some way to explain the greater numbers of men achieving distinctions of various kinds for which a high IQ is required, such as chess Grandmasters, Fields medallists for mathematics, Nobel prize winners and the like’ (p. 519; emphasis added).

Despite the increasing proportion of women in intellectually demanding professions, they are still under-represented at the top level in science (Long 2001; Xie & Shauman 2003; Ceci & Williams 2007) and engineering (Long 2001). There are various possible explanations for this apart from innate biological differences (for critiques and negative findings see Kerkman et al. 2000; Spelke 2005; Lachance & Mazzocco 2006). Socialization and different interests (Benbow et al. 2000; Ayalon 2003), gender roles (Massa et al. 2005), gatekeeper effects (Davidson & Cooper 1992; Steele 1997; Huffman & Torres 2002), cultural differences (Andreescu et al. 2008) and higher participation rates of men (Charness & Gerchak 1996; Chabris & Glickman 2006) have all been proposed. Here, we show that in chess, an intellectually demanding activity where men dominate at the top level, the difference in the performance of the best men and women is largely accounted for by the difference that would be expected, given the much greater number of men who participate. Despite the clear superiority of the top male players, there is, in reality, very little performance gap in favour of men for non-statistical theories to explain.

Before considering cultural or biological explanations for better performance by the top performers in one of two groups of different size, a simple statistical explanation must be considered. Even if two groups have the same average (mean) and variability (s.d.), the highest performing individuals are more likely to come from the larger group. The greater the difference in size between the two groups, the greater is the difference to be expected between the top performers in the two groups. Nothing about underlying differences between the groups can be concluded from the preponderance of members of the larger group at the far ends of the distribution until one can show that this preponderance is greater than would be expected on statistical sampling grounds.

It is difficult to quantify how participation rates influence the number of outstanding men and women in fields such as science and engineering because both achievement and participation rates are difficult to measure. But it is straightforward in chess because there is an objective measure of achievement and the number of male and female participants is known. Chess has an
We applied this method to the population of all German players recorded by the German chess federation (Deutscher Schachbund). With over 3000 rated tournaments in a year, the German chess federation is one of the largest and the best organized national chess federations in the world. Given that almost all German tournaments are rated, including events such as club championships, all competitive and most hobby players in Germany can be found on the rating list. The rating itself is based on the same assumptions as the Elo rating used by the international chess federation. The two correlate highly ($r=0.93$).

We considered the players in the list published in April 2007. A small number of the best male and female chess players from all over the world participate in tournaments rated by the German chess federation and dominate the top of the rating list. We have excluded all foreign players from the analyses so that our conclusions about the expected performance of the best male and female players, based on the total number and performance of male and female German players, are only applied to German players. Figure 1 shows the distribution of ratings for the German chess population. The distribution is approximately normal with mean of 1461 and s.d. of 342. Rated men (113 386) greatly outnumber rated women (7013); that is, there are 16 male chess players for every woman.

3. RESULTS

Figure 2 shows the real difference in rating for each of the top 100 pairs of male and female players, and the difference to be expected for each pair given the much larger number of male players. The expected superiority of male players varies from approximately 270 Elo points for the best male player to approximately 440 Elo points for the 100th. Figure 2 shows that, in fact, the top three women are better than would be expected. The next 70 pairs show a small but consistent advantage for men—their superiority over the corresponding female player is a little greater than would be expected purely from the relative numbers of male and female players. From approximately the 80th pair the advantage shifts. The female players are slightly better than would be expected. Averaged over the 100 top players, the expected male superiority is 341 Elo points and the real one is 353 points. Therefore 96 per cent of the observed difference between male and female players can be attributed to a simple statistical fact—the extreme values from a large sample are likely to be bigger than those from a small one.

4. DISCUSSION

Chess has long been renowned as the intellectual activity par excellence (Newell et al. 1958) and male dominance at chess is frequently cited as an example of innate male intellectual superiority (e.g. Howard 2005; Irwing & Lynn 2005). The reason seems obvious—the best male players are indisputably better than the best female players. For example: not a single woman has been world champion; only 1 per cent of Grandmasters, the best players in the world, are female; and there is only one woman among the best 100 players in the world. When considering such a seemingly convincing example of real world male superiority, one can easily forget to consider the great disparity in the number of participants and the statistical consequences of this for the probable gender of the best players.
This was the case when the chess portal ChessBase asked some of the best female players to explain male dominance in chess (Ahmadov 2007). None of the interviewed women even mentioned the greatly differing participation rates and its consequences on the probable gender of the top performers. Similarly, at a recent gathering of more than 20 experts on gender difference to discuss the reasons for the paucity of women at the top of science, a broad range of reasons was discussed, but there was no mention of participation rates (Ceci & Williams 2007).

One way to avoid the conclusion based on participation rates would be to argue that the base participation rate for women in this study underestimates the real participation rate. It is possible that there is a self-selection process based on the innate biological differences in intellectual abilities, and that the effects of this self-selection are already observable in the rating list we used. Women may be inferior in the intellectual abilities that are important for successful chess playing. This innate disadvantage may lead women to give up on chess in greater numbers than more successful men. The small number of women is then a consequence of their greater drop-out, which in turn is produced by their innate lack of the intellectual abilities required to succeed at chess. Differential participation rates may explain the discrepancy at the top, but the difference is itself a direct product of innate differences in intellectual abilities.

This argument sounds reasonable but it rests on a controversial assumption. It requires that there should be innate differences between men and women in the intellectual abilities required for success at chess. The topic of gender differences in cognitive abilities is a hotly debated one, which lacks conclusive evidence (for example, Geary 1998; Kimura 1999; Kerkman et al. 2000; Pinker 2002; Spelke 2005; Summers 2005; Lachance & Mazzocco 2006; Ceci & Williams 2007). Even if such differences exist, it is unclear which, if any, intellectual abilities are associated with chess skill (for a recent review, see Bilalić et al. 2007). Whatever the final resolution of these debates, there is little empirical evidence to support the hypothesis of differential drop-out rates between male and females. A recent study of 647 young chess players, matched for initial skill, age and initial activity found that drop-out rates for boys and girls were similar (Chabris & Glickman 2006). Our study does not deal directly with the reasons why there are so few women in competitive chess. These may have to do with selective drop-out before tournament play starts in the early stages of learning to play chess. We can speculate about the reasons for low participation rates of women in competitive intellectual endeavours (as is often done, e.g. Steele 1997; Benbow et al. 2000; Kerkman et al. 2000; Massa et al. 2005; Spelke 2005; Summers 2005; Lachance & Mazzocco 2006; Andreescu et al. 2008) but empirical evidence is scarce.

This study demonstrates that the great discrepancy in the top performance of male and female chess players can be largely attributed to a simple statistical fact—more extreme values are found in larger populations. Once participation rates of men and women are controlled for, there is little left for biological, environmental, cultural or other factors to explain. This simple statistical fact is often overlooked by both laypeople and experts. In other domains such as science and engineering, where the predominance of men at the top is offered as evidence of the biological superiority of men, large differences between the number of women and men engaged in these activities are evident (Long 2001; Xie & Shauman 2003). In these areas of life, it is not possible to estimate the performance of the top women and men and their participation rates as precisely as it is in chess. But until
which is equivalent to the expected value (or mean) of a sample of size $n$ respectively, in our sample. Now defining a simple approximation valid for large $X_n$ is all ratings from the German chess federations for both identically distributed variables; in our case, the sample of terms of the form $E_{m,1}$—the best performer from a sample of size $m$—where the indexing variable takes values between $n-k$ and $n$.

Suppose that $X_i \sim N(0,1)$ are drawn from a standard normal distribution. If the sample size is large, then the expected value of the best performer increases linearly with the logarithm of sample size, i.e. $E_{m,1} \approx c_1 + c_2 \ln m$, where $c_1 = 1.25$ and $c_2 = 0.287$ (figure 3).

If $n-k$ is large, then each value taken by $m$ in the previous sum will be large. Hence, we may apply the log-linear approximation to each occurrence of $E_{m,1}$, obtaining

$$E_{n,k+1} \approx c_1 + c_2 \frac{n!}{k!(n-k-1)!} \sum_{m=n-k}^{n} (-1)^{k+m-n} \binom{k}{m} \frac{\ln m}{m}$$

Here, $f(x) = \ln x/x$ and $\varphi(x) = \ln x - \ln(x-1)$ denotes the backward difference operator (Abramowitz & Stegun 1972). Furthermore, the backward difference operator is a close approximation of the differential operator, and hence

$$\nabla^k f(x) = \frac{d^k}{dx^k} f(x) = (-1)^k k! \frac{\ln x - H(k)}{x^{k+1}}$$

where $H(k) = \sum_{i=1}^{k} i^{-1}$ denotes $k$th harmonic number and we define $H(0) = 0$. Substituting, we find that the expected value of the $k$th best performer is given by

$$E_{n,k} = c_1 + c_2 \frac{n!}{(n-k)!} \ln(n - H(k - 1))$$

This holds if $X_i \sim N(0,1)$ are drawn from a standard normal distribution. Moving to a general normal distribution, if $X_i \sim N(\mu, \sigma^2) = \mu + \sigma N(0,1)$, we find

$$E_{n,k} = (\mu + c_1 \sigma + c_2 \frac{n!}{(n-k)!} \ln(n - H(k - 1))$$

Given a distribution with known mean $\mu$ and s.d. $\sigma$, this final formula defines the expected value of the $k$th highest value within a sample of size $n$, valid provided $n$ is large and $k$ is relatively small. As such, it affords us a method for estimating the expected rating of a range of top players from the German chess data for each gender; indeed, we use the formula to calculate the expected ratings of the top 100 male and female players using the mean and s.d. of the population (the German chess data), in turn allowing us to determine the expected difference in rating between those players.

REFERENCES


