Running head: Sex differences in competitive chess performance

Sex Differences in Intellectual Performance: Analysis of a Large Cohort of Competitive Chess Players

Christopher F. Chabris
Harvard University

Mark E. Glickman
Boston University School of Public Health

Address correspondence to: Mark E. Glickman
Center for Health Quality, Outcomes & Economics Research
Edith Nourse Rogers Memorial Hospital (152), Bldg 70
200 Springs Road
Bedford, MA 01730 USA
mg@bu.edu
tel: 781-687-2875

3980 words

Abstract

Only 1% of the world's chess grandmasters are women. This underrepresentation is unlikely to be caused by discrimination, since chess ratings objectively reflect competitive results. Using data on the ratings of 250,000 tournament players over 13 years, we investigate several potential explanations for the male domination of elite chess. We find that: (1) the ratings of men are higher on average than those of women, but no more variable; (2) matched boys and girls improve and drop out at equal rates, but boys begin chess competition in greater numbers and at higher performance levels than girls; (3) in locales where at least 50% of the new young players are girls, their initial ratings are not lower than those of boys. We conclude that the greater number of men at the highest levels in chess can be explained by the greater number of boys who enter chess at the lowest levels.

Sex Differences in Intellectual Performance:

Analysis of a Large Cohort of Competitive Chess Players

The game of chess has been studied by computer scientists and cognitive psychologists as a model arena of human intellectual performance. Research on computer chess has culminated in programs that can defeat the best human players (e.g., Hsu, 2002), while research on chess masters has yielded seminal discoveries such as the chunk structure of short-term memory (Chase & Simon, 1973) and has contributed to debates on the importance of pattern recognition and deliberate thought in expertise (Gobet & Simon, 1996; Chabris & Hearst, 2003; Burns, 2004). But one of the most striking facts about chess competition has received little study: the dramatic lack of women among the game's elite performers. None of the official world champions was a woman, no champion of a major country is a woman, and as of January 2004 only 9 of the world's 894 chess grandmasters—1%—were women (according to data in Howard, 2005).

Analyzing possible explanations for the underrepresentation of women among the chess elite may help us understand the underrepresentation of women at the highest levels in other fields, such as tenured professorships in mathematics, science, and engineering. It has been suggested (e.g., Summers, 2005; Pinker, 2005) that differences between men and women in the distribution of cognitive abilities required for success in these fields can partly account for the disparity (the *ability distribution hypothesis*). In particular, men and women may differ in mean performance levels and/or variability of performance; evidence suggests that in cognitive abilities, both types of differences are found (Halpern, 2000; Hedges & Nowell, 1995).

However, the possibility of "old boys networks" of men as gatekeepers to high positions in these fields, coupled with the subjective nature of assessing achievement, makes it difficult to distinguish between an objective lack of achievement or credentials and discrimination by the existing social system as causes. In chess, neither of these conditions obtain; in particular, the rating system invented by Elo (1986) objectively measures individual skill based only on results of tournament games. The U.S. Chess Federation (USCF) applies this system to rate tens of thousands of players who participate in events that are open to all. Therefore, the overrepresentation of men at the highest levels in chess is, at first glance, more consistent with an ability distribution hypothesis than with a social-system account. (Note that a difference in mean, variance, or both could explain the observed differences at the upper tail of the distribution.)

However, other explanations are possible, such as differential dropout over time. Men and women may start out with equal endowments of the abilities necessary for an endeavor, but women may be less likely than men to intensively study, practice, or devote obsessive amounts of time to it (the *differential dropout hypothesis*). Indeed, some argue that it is precisely the amount of "deliberate practice" that predicts success in fields like chess (Ericsson et al., 1993). In chess, this would mean that more *potential* female than male grandmasters leave organized competition, resulting in an imbalance at the top levels.

Anyone who visits an open chess tournament will be struck less by the lack of women at the top of the results table than by their near absence at all levels. Only 9.7% of all USCF-rated games in 2004 were played by women. It is possible that the lack of women at the top is an artifact of their lower overall participation rate (Charness & Gerchak, 1996): Even if men and women have the same underlying ability distribution, a larger number of top-rated players will be men if the overall number of men competing is greater (the *participation rate hypothesis*). That is, if fewer women than men even begin to participate in organized competition, dropout rates (and cognitive endowments) could be equal, but women would still be relatively absent at the top.

Here we ask whether these three hypotheses explain the enormous imbalance between men and women among the best chess players. Previous research on sex differences in chess performance (Charness & Gerchak, 1996; Howard, 2005) has only considered players at the top end of the rating spectrum. In this study, we analyze the annual ratings of nearly all of the chess players who participated in USCF-rated games over 13 years, from 1992 through 2004. This is the broadest and largest sample of chess performance data ever analyzed, and one of the best datasets on sex differences in intellectual performance in any domain.

General Method

The data for our study included rating information on all USCF members that have both birth date and sex recorded in the USCF database, a total population of 256,741 tournament players. Table 1 shows the sex distribution by players' ages; 10.6% of the sample is female.

Insert Table 1 Here

For each player, we recorded birth date, sex, most recent zip code (if available), and year-end rating on each annual rating list from 1992 through 2004. A rating only appears on the annual list if the player played at least one rated game that year. For the years of our study, most players have some missing rating information because either they were inactive during one or more years, or they started playing in USCF tournaments after 1992. We also recorded the number of tournament games played per year by each player, and whether the rating was "provisional" or "established." Provisional ratings are based on fewer games than established ratings, and are generally less reliable measures of playing strength. Once a player's rating becomes established, it remains established.

A player's USCF rating is an estimate of his/her current playing strength on a scale that ranges generally from 100 to 3000, with higher ratings associated with better playing ability. The principles underlying the rating computations are explained by Elo (1986) and Glickman (1995).[1] Average tournament players are usually rated between 1400 and 1600, chess masters are rated above 2200, and world-class players tend to be rated above 2500. USCF ratings are essentially estimates of merit parameters from the Bradley and Terry (1952) model for paired comparisons, using an approximately Bayesian filtering algorithm to update ratings over time (Glickman, 1999). While ratings are only estimates of unknown parameters and are therefore subject to variability, they can be treated as data in statistical modeling, recognizing that the extra uncertainty may lead to conservative inferences.

Cross-sectional Analyses of Sex Differences

We begin by asking whether a difference in mean chess ability exists between male and female tournament competitors, whether any such difference depends on age, and whether it has changed over time. We calculated the mean difference in ratings between males and females for each year of the study, restricting the sample to players with established ratings who played at least one rated game in the given year (see Figure 1, left panel).

Insert Figure 1 Here

On average, mean ratings are higher for males by 450–500 points, a very large difference: The expected outcome of a game between opponents having this rating difference is 0.93–0.95 (counting a win by the higher-rated player as 1, a loss as 0, and a draw as 0.5). We adjusted these simple sex differences by incorporating covariates in linear regression models. For rating lists from 1995 through 2004, we examined the subsample of players with established ratings in the given year, and modeled current rating as a function of sex, current age, number of games played in the current year, and number of games played in the previous three years. Prior to model fitting, the latter three explanatory variables were discretized into categories because we did not want to assume *a priori* that they related to ratings via a simple parametric function (Elo, 1986; Charness et al., 1996). The breakpoints for the three numerical predictor variables were chosen by fitting regression trees (Breiman et al., 1984) to the 2004 data, regressing current rating on each variable separately, resulting in the following categories:

- Age (years): < 12.8, 12.8–17.5, 17.5–19.5, ≥ 19.5
- Number of games in current year: 0–9, 10–50, ≥ 51
- Number of games in prior three years: 0–74, 75–151, ≥ 151

The choice of these breakpoints was supported by similar analyses of data from the earlier years.

A regression of current rating on sex, adjusted by the full interaction of the three other predictor variables, resulted in average adjusted sex differences shown in the right panel of Figure 1. While accounting for current and past frequency of play (both of which were related positively to rating) and age decreased the sex difference by a factor of about 3, the mean difference each year was still a highly significant 150–200 points (corresponding to an expected game outcome of 0.70–0.76). Similar analyses using alternative breakpoints, or keeping the variables untransformed, yielded similar results.

We also examined sex differences for the model where sex and the three other variables were fully interacted. Fitting this model is analogous to examining sex differences in rating separately within strata formed by all combinations of the categorized variables (age, number of current games, and number of games in the prior three years). We performed this analysis separately for each year from 1995 to 2004. In strata with more than 25 females, males always had higher ratings (and significantly so, with very few exceptions). While the mean difference dropped as low as 62.8 (in 2002, for 12.8–17.5 year old players who played the fewest games in the current year and prior years), there was no clear pattern to the effect of age and playing frequency on sex differences.

Sex Differences in Rating Variation

Regardless of any difference in mean ratings, the disproportionate number of men at the top in chess could result from their abilities being more variable in the general population. This larger male variation would also imply a larger proportion of men at the bottom. This argument is consistent with data on cognitive test scores (Hedges & Nowell, 1995) and it has been offered to explain the high male:female faculty ratios in academic disciplines (e.g., Summers, 2005; Pinker, 2005). Because only the upper tail of the distribution of chess players self-select to compete in tournaments, this hypothesis would not necessarily predict that the lowest rated players should be male, but the rating variation should still be larger for men than for women.

To examine this, we computed the standard deviation of ratings for males and females, stratified by the age groupings used in Table 1 (collapsing ages 65–95 into one group), for each year in our data. The standard deviations by gender/age/year strata range from about 250 to 500. The left panel of Figure 2 displays the female-to-male ratios of standard deviations. Generally, the ratios are greater than 1, most particularly for players between 25 and 55, meaning that female rating variation is typically larger than that for males. Only in the extreme age groups are the ratios close to 1. The data clearly do not support greater male variation in ratings.

Sex Differences in Longitudinal Rating Changes

The consistently higher mean male rating observed so far, with no clear sex difference in variability, could be explained by girls beginning to play on an equal footing with boys, but improving more slowly or dropping out in greater numbers, so that those women who remain in the rating system have a lower mean rating than the men. To test this hypothesis, we performed a case-control study by creating a subsample of our dataset in which each female was initially matched as closely as possible to a male, and then following these pairs over time. This analysis focused on players aged 5–25 years old in 1995 who had established ratings. For each player, we recorded four variables: 1995 year-end rating, age, number of games played in 1995, and number of games played in the previous three years. Younger players were used for this analysis because we wanted to examine a group who were recent entrants to tournament chess. We then formed male/female pairs via caliper matching (Cochran & Rubin, 1973) on these four variables, with a common caliper size of 0.15. That is, within each pair, the male and female values on each variable differed by no more than 0.15 standard deviations of the overall distribution of that variable. This process resulted in 647 matched pairs, which we tracked for 10 years. The right panel of Figure 2 shows the mean rating difference with 95% confidence intervals for each year with 10 or more pairs; the difference does not significantly deviate from 0. Similar analyses with different starting years, smaller calipers, and different age ranges, yielded the same conclusions.

On the basis of research on cognitive sex differences (reviewed by Kimura, 1999), one might suggest that males should overtake females in chess performance only around puberty. That is, if chess skill relies in part on visual-spatial ability (Robbins et al., 1996; Frydman & Lynn, 1992; but see Waters et al., 2002), and if this ability is influenced by testosterone, then males might benefit from the increase in androgens during the early teen years. Differences in spatial task performance have been observed long before puberty (e.g., before age 5 by Levine et al., 1999; ages 8–9 by Levine et al., 2005), but the magnitude of the male advantage tends to increase during the teenage years (see Table 4 of Voyer, Voyer, & Bryden, 1995). If this phenomenon applied to chess performance, it would appear in the rating comparison of males and females during their early teens who had similar pre-teen ratings. We investigated this in our matched pair sample by examining the within-pair rating difference at the end of 1998 as a function of the players' ages in 1995, and found that these differences were evenly distributed around 0, suggesting that males on average were not overtaking females during the early teen years. Analyzing data from other years yields the same conclusion.

Interestingly, the attrition of males and females in our matched pairs group is remarkably similar. The percentage of females among the active players in any given year is no less than 47.7% (in 1996) and no higher than 55.1% (in 2004). Not surprisingly, attrition among females in the group of all established players in 1995 aged 5–25 from which our matched sample was drawn is greater; on average, females continue playing beyond 1995 for 1.74 years until becoming inactive, compared to 1.95 years for males ($p < .0001$), consistent with the disproportionately fewer women at the higher age groups (Table 1). The greater female attrition could result simply from lower-rated players, regardless of sex, tending to become inactive more quickly because they lose interest, become discouraged, etc. In fact, this hypothesis is supported by our data. We performed a Poisson regression of the number of years until a player first becomes inactive on 1995 rating, age group (using the CART-determined categories), and sex. Both 1995 rating and age group were highly significant (lower rating and older ages predicted attrition), but sex was not ($p = .54$, likelihood ratio test).

Thus, while men have a higher average rating than women, matched samples of boys and girls play with the same frequency, are equally likely to drop out, and improve their ratings at the same rate, without diverging at puberty. The tendency for women to drop out more frequently overall appears to be related to age and playing strength, but not sex.

Sex Differences in Initial Ratings of New Tournament Players

If males have higher mean ratings, but there are no sex differences in rating change over time, then they must start out in tournaments with higher ratings than females. To confirm this, we examined for each year from 1998 through 2004 the set of players ages 6–12 who had established ratings at year-end, and who did not have a rating in any year before the previous one. We excluded older players because they could have been playing chess frequently before beginning tournaments. We chose to restrict our analysis to 1998 and later because it was difficult to determine whether a player had competed in tournaments previously (specifically, prior to the start of our cohort in 1992). On average, the sex difference in ratings for these groups was 110–200 points in favor of the males, and was always highly significant ($p \ll .0001$). Linearly adjusting for age (whose effect was 20–45 rating points per year for this age group) did not change the significance or magnitude of the sex difference.

Finally, we address the participation rate hypothesis. If the number of boys in the general population that play chess is substantially larger than the number of girls, the best ones ultimately becoming USCF members and playing competitively, then it follows statistically that the average boys' ratings will be higher than girls' ratings (among competitive players) even if the distribution of abilities in the general population is the same (Charness & Gerchak, 1996; Glickman & Chabris, 1996). In fact, far fewer girls than boys enter competitive chess, which suggests that the general population of chess-playing girls is much smaller than that of boys. External factors like the relative lack of women among the world's top players to serve as role models, and the prospect of playing a game dominated by boys, may be discouraging to girls (or their parents), either directly reducing their likelihood of learning how to play in the first place, or indirectly reducing their initial performances in competitive play via a "test anxiety" or "stereotype threat" mechanism (Steele, 1997). Thus, it is possible that, on average, girls have the chess-relevant cognitive abilities, but the larger number of boys playing chess leads to significantly higher male ratings in the USCF population.

To test this, we searched for locales in which the female participation rate is high, and determined the extent of sex differences in initial ratings. We assume that in places where girls play competitive chess as commonly as boys, the social factors ordinarily discouraging girls from playing chess may be minimal. If this is true, then the participation rate hypothesis predicts no difference between boys' and girls' ratings.

We examined a subsample of the players between 6 and 12 years old who had an established rating at year-end, and did not have a rating in any year before the previous one. We included only players for whom we had 2004 zip code information, and excluded years prior to 2002 because the 2004 zip codes become less reliable indicators of place of residence as we move farther back in time. For these three years, we measured sex differences in rating within zip code (excluding zip codes with fewer than 10 players).[2] Figure 3 shows the mean rating difference as a function of the proportion of females in the zip code area (truncated at 0.2 from below).

Insert Figure 3 Here

Boys generally have higher ratings than girls, particularly on the left side of the plot where zip codes are male-dominated. However, in the four zip codes with at least 50% girls (areas in Oakland, CA, Bakersfield, CA, Lexington, KY, and Pierre, SD), boys no longer have higher ratings. In Oakland, with the greatest proportion (68%) of girls, the average rating of girls is higher than that of boys, though not significantly so. Combining all zip code areas where the proportion of girls is at least 50%, the sex difference is only 35.2 points in favor of males, which is not significant (p = .589). This conclusion also holds in an age-adjusted analysis, where the sex difference is 40.8 points (p = .532).

The fairly constant mean male advantage all the way until the 50% female participation rate suggests that a threshold effect may be operating, such that factors limiting girls' performance levels depend on their being in the minority, but not on the relative size of the male majority (in other words, 50% girls may constitute a "critical mass"). Further study of this effect is warranted, perhaps by searching for chess rating data from larger social-geographic regions, such as nations, across which the female participation rate varies, especially cases (if any) in which competitive chess is significantly *more* popular among girls than boys.

Discussion

This study analyzed 13 years of chess rating data to discover the reason for the vast overrepresentation of men among elite players, exemplified by the 99:1 male:female ratio of international grandmasters. We found that men were rated an average of 150–200 points higher than women on the Elo (1986) scale, even after controlling for age and frequency of play, a highly significant male advantage. A longitudinal analysis of matched male-female pairs showed that girls and boys of equal strength do not diverge in playing ability or frequency; instead, boys and girls enter competitive chess with different average ability levels, and this difference propagates throughout the rating pool. However, this initial difference is not found in locales where boys and girls enter the rating system in equal proportions. Taken together, our results support the hypothesis that there are far fewer women than men at the highest level in chess because fewer women enter competitive chess at the lowest level (which is consistent with men and women having equal chess-relevant cognitive abilities).

Consistent with our overall findings, Charness and Gerchak (1996) showed that the rating difference between the world's top male and female players in the mid-1990s (Gary Kasparov of Russia and Judit Polgar of Hungary) could be explained by men and women having equal means but different participation rates.[3] By contrast, Howard (2005) argued that social factors were unlikely to explain the sex difference observed among top international players, because the difference has persisted over the past 30 years despite presumed worldwide increases in the opportunity and encouragement for women to enter competitive chess. However, Howard's argument does not take account of the vastly different numbers of men and women in even the current international rating lists; moreover, his own data show that since the international rating rules were changed so that the minimum rating was the same for men and women, the average male rating has drifted down towards the average female rating (while the gap between the top 100 men and top 100 women has remained constant).

We have shown that sex differences in participation rates in the population of rated chess players can explain the vast overrepresentation of men among the game's elite. This is somewhat surprising given the relative objectivity of chess skill measurement and the lack of subjective judgment in determining competitive achievements. Thus, significant male-female differences can arise in elite performance even in the absence of gatekeeping mechanisms and advancement standards controlled or developed by men. Beyond the domain of competitive chess, our results show how male-female differences in factors other than cognitive abilities may account for sex differences in observed performance. Understanding the causes of these factors—in our case, why more boys than girls enter competitive chess—is a challenge in itself.

References

Bradley R.A., & Terry, M.E. (1952). The rank analysis of incomplete block designs. 1. The method of paired comparisons. *Biometrika*, *39*, 324–345.

Breiman, L., Friedman, J.H., Olshen, R.A. , & Stone, C.J. (1984). *Classification and regression trees.* Monterey, CA: Wadsworth and Brooks/Cole.

Burns, B.D. (2004). The effects of speed on skilled chess performance. *Psychological Science, 15,* 442–447.

Chabris, C.F., & Hearst, E.S. (2003). Visualization, pattern recognition, and forward search: Effects of playing speed and sight of the position on grandmaster chess errors. *Cognitive Science, 27,* 637–648.

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.

Charness, N., Krampe, R., & Mayr, U. (1996). The role of practice and coaching in entrepreneurial skill domains: An international comparison of life-span chess skill acquisition. In Ericsson, K.A. (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports and games* (pp. 51–80). Mahwah, NJ: Erlbaum.

Chase, W.G., & Simon, H.A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55–81.

Cochran, W., & Rubin, D.B. (1973). Controlling bias in observational studies. *Sankhya, 35,* 417–446.

Elo, A.E. (1986). *The rating of chessplayers, past and present* (2nd ed.). New York: Arco.

Ericsson, K.A., Krampe, R.T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406.

Frydman, M., & Lynn, R. (1992). The general intelligence and spatial abilities of gifted young Belgian chess players. *British Journal of Psychology, 83,* 233–235.

Glickman, M.E., & Chabris, C.F. (1996). Using chess ratings as data in psychological research. Unpublished manuscript. [www.wjh.harvard.edu/~cfc/Glickman1996.pdf]

Glickman, M.E. (1995). Chess rating systems. *American Chess Journal, 3,* 59–101.

Glickman, M.E. (1999). Parameter estimation in large dynamic paired comparison experiments. *Applied Statistics,* 48, 377–394.

Gobet, F., & Simon, H.A. (1996). The roles of recognition processes and look-ahead search in time-constrained expert problem solving: Evidence from grandmaster-level chess. *Psychological Science, 7,* 52–55.

Halpern, D. (2000). *Sex differences in cognitive abilities* (3rd ed.). Mahwah, NJ: Erlbaum.

Hedges, L.V. and Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269,* 41–45.

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

Hsu, F-H. (2002). *Behind Deep Blue: Building the computer that defeated the world chess champion.* Princeton, NJ: Princeton University Press.

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

Levine, S.C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology, 35,* 940–949.

Levine, S.C., Vasilyeva, M., Lourenco, S.F., Newcombe, N.S., & Huttenlocher, J. (2005). Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science, 16,* 841–845.

Pinker, S. (2005). The science of difference. *The New Republic,* 14 February.

Robbins, T.W., Anderson, E.J., Barker, D.R., Bradley, A.C., Fearneyhough, C., Henson, R., Hudson, S.R., & Baddeley, A.D. (1996). Working memory in chess. *Memory and Cognition, 24,* 83–93.

Steele, C.M. (1997). A threat in the air: How stereotypes shape the intellectual identities and performance of women and African-Americans. *American Psychologist, 52,* 613–629.

Summers, L.H. (2005). Remarks at NBER conference on diversifying the science and engineering workforce. Cambridge, MA, 14 January. [www.president.harvard.edu/speeches/2005/nber.html]

Voyer, D., Voyer, S., & Bryden, M.P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117,* 250–270.

Waters, A., Gobet, F., & Leyden, G. (2002). Visuo-spatial abilities in chess players. *British Journal of Psychology, 30,* 303–311.

Author Note

Christopher F. Chabris, Department of Psychology, Harvard University. Mark E. Glickman, Department of Health Services, Boston University School of Public Health.

We thank Carole K. Hooven and Joel Reisman for helpful discussions, and especially Michael Nolan of the USCF for generously providing the data analyzed here.

Address correspondence to: Mark E. Glickman, Center for Health Quality, Outcomes & Economics Research, Edith Nourse Rogers Memorial Hospital (152), Bldg 70, 200 Springs Road, Bedford, MA 01730 USA; Email: mg@bu.edu.

Endnotes

1. For complete details of the USCF rating system, see www.uschess.org.

2. In other analyses not reported here, we grouped players by larger geographic regions as functions of zip code (metropolitan statistical area, FIPS code, and urban area code), but these groupings combined locales with high proportions of females with many locales with low proportions, leaving little variation in female proportion across the resulting regions.

3. Note, however, that an analysis of extremes in a distribution is a very low-power method of inferring differences in the sample means (Glickman & Chabris, 1996). Here we have analyzed the entire distribution, rather than just the extreme performers.

Table 1. *Distribution of females within 10-year age groups, and distribution of ages in study population.* The highest concentration of female players is in the younger age groups, which comprise the majority of tournament players. By the age of 35, only about 2% of tournament players are women.

Age (years)	Female percent in age group	Percent of study population
5–15	17.0	26.4
15–25	11.9	43.5
25–35	4.9	11.5
35–45	2.2	6.4
45–55	2.0	6.8
55–65	2.2	3.3
65–75	2.0	1.3
75–85	2.1	0.7
85–95	2.5	0.1

Figure 1. Left: Mean male–female rating difference by year, with 95% confidence intervals. Right: Regression-adjusted estimates of male–female rating difference by year, with 95% confidence intervals.

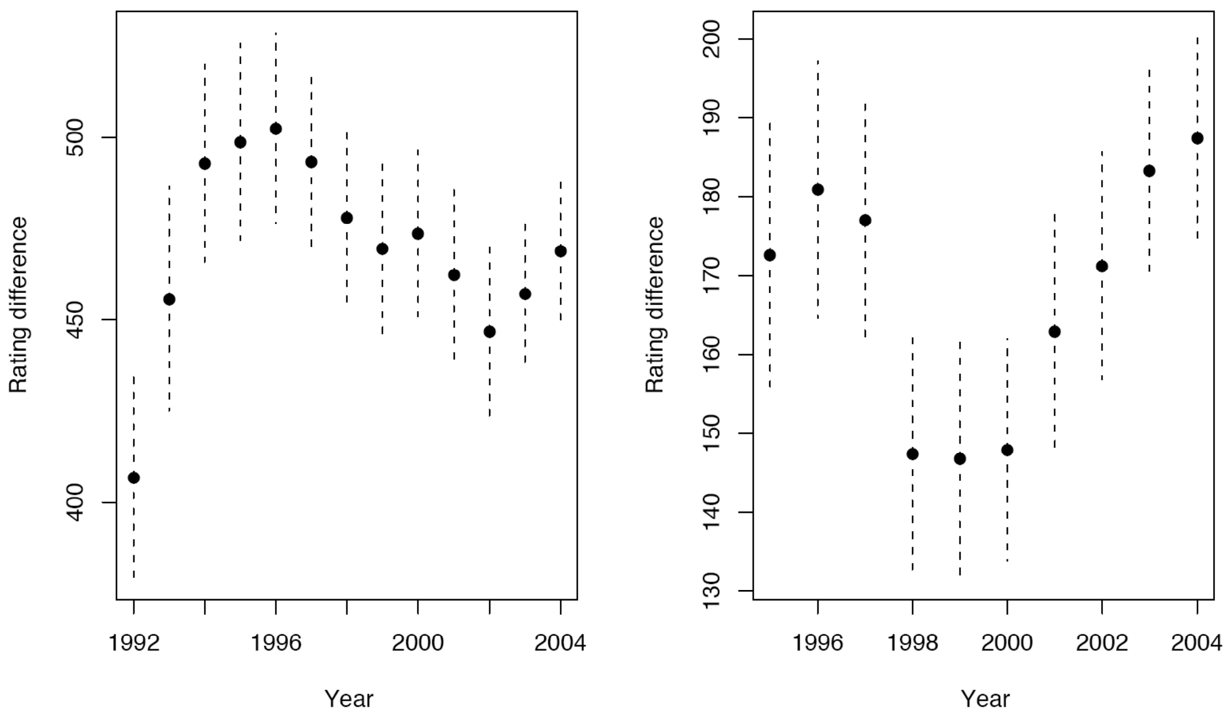

Figure 2. Left: Ratio of standard deviations (female/male) of established ratings by age grouping for years 1992−2004. Each open circle represents the ratio for that age group in one year, and the solid circles are the mean ratios across the 13 years. Right: Mean and 95% confidence intervals for male–female rating differences by year for the matched pairs analysis. (The wider confidence intervals for later years reflect the attrition of male/female pairs due to players becoming inactive.)

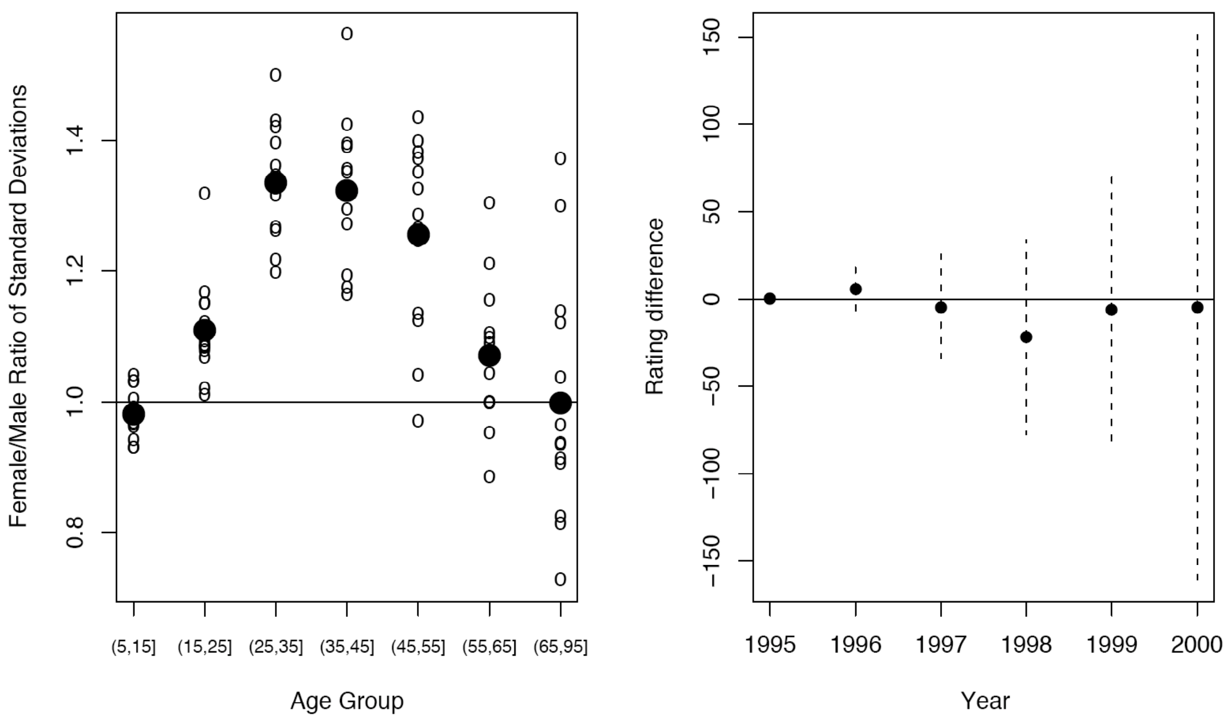

Figure 3. Mean male–female rating difference for young established players in 2002−2004 by proportion of females in zip code area, with 95% confidence intervals.