A Computational Aesthetics Assessment of Chess Playing Quality Between the Genders

Azlan Iqbal[1+], Siva Nagappan[2]

[1] College of Computer Science and Information Technology, Universiti Tenaga Nasional, Putrajaya Campus, Selangor, Malaysia.

[2] College of Engineering, Universiti Tenaga Nasional, Putrajaya Campus, Selangor, Malaysia.

Abstract. Aesthetics or beauty in chess is a quality appreciated by most players. However, there is scant research on the differences of aesthetic perception between the genders, especially given the lower participation of females in this domain. Using an experimentally-validated computational aesthetics model for chess, we evaluated a fair selection of winning chess move sequences taken from games played between women and men. Contrary to previous research that was not as thorough, we found no statistically significant difference in the aesthetic quality of those sequences between the groups. The results suggest that aesthetic ability, perception and appreciation in the game are likely not affected by gender. This also implies that training methods and promotion of the game to girls or young women have less, if any, basis for being any different from those that pertain to boys or men. Furthermore, the arguably absolute lack of participation of women in the sub-domain of chess problem composition – in which aesthetics plays an even more significant role – likely has little, if anything, to do with innate capability unless otherwise demonstrated.

Keywords: chess, aesthetics, gender, perception, artificial intelligence.

1. Introduction & Background

Aesthetics or beauty in chess tends to be found in move sequences in the game which demonstrate particular properties such as expediency (i.e. effectiveness), disguise (e.g. an unclear initial move), material sacrifice, correctness, preparation (e.g. strategic play preceding the move combination), paradox (e.g. going against rules of thumb), unity (cooperation between the pieces toward the goal) and originality [1-7]. Aesthetics is therefore generally appreciated in the game and one of the motivations for players to pursue it either professionally or as an intellectual pastime. In the sub-domain of chess problem composition, aesthetics plays an even more significant role. Aesthetic perception (relating also to the capability to produce beauty and the appreciation of it) differences in the game between men and women, however, is not well understood. This is despite the significantly lower participation of females in the sport. Learning of such differences, if any, would be beneficial in terms of training and promotion methods used to introduce children to the game [8-9]. It would also contribute to the literature of male versus female psychology in this domain and perhaps even strategic board games in general.

1.1. Previous work & objectives

Earlier research into gauging aesthetic differences of perception in chess between the genders showed that the games played between males and the games played between females had a statistically significant difference, on average, in favour of those between males [10]. However, the sequences analyzed (using the same model mentioned in section 2.1 of *this* paper) were limited to only three-move mate sequences. In addition, playing strength (i.e. player Elo rating at the time of play) was not taken into account. The difference in aesthetic quality of play could have been due simply to playing strength because men, on

---

+ Corresponding author. Tel.*: +603 89212334; *fax*: +60389212334.
  E-mail address: azlan@uniten.edu.my

average, tend to have higher ratings than women despite some controversy in this area of research as well [11-17]. Our objectives in the present research were therefore; 1) to perform a more thorough examination of the aesthetic quality of chess sequences taken from games between men and women using artificial intelligence (AI) as in the previous work; 2) to determine if there were any significant differences and; 3) to use the results in order to derive recommendations for training and promotion methods to females in order to improve their participation in the sport.

2. Methodology

2.1. The Computational Aesthetics Model

Similar to previous work on this topic [10] the computer program, Chesthetica, was used to automatically analyze the aesthetics in the game of chess. Chesthetica incorporates an experimentally-validated computational aesthetics model that is able to evaluate aesthetics in the game in a way that correlates positively and well with domain-competent human assessment [18]. This has its advantages over manual human expert assessment in terms of cost, reliability and consistency. The model uses formalizations of seven aesthetic principles and ten chess themes combined with a stochastic approach to produce a numerical aesthetic score (precise to three decimal places) for a chess move sequence. As an unintended consequence, this score may vary very slightly each time a position and its sequence is evaluated; not unlike with a human expert taking a 'second look'. For the purposes of this research, just one cycle of evaluation was used for each position and the score, whatever it may be, was taken as final.

Human experts would also have evaluated the positions just once as doing so multiple times would have been unnecessarily burdensome and confusing to them. Chesthetica is not an open-source program or freely available for download but further details including how to calculate the aesthetics of a move sequence manually are provided in the aforementioned reference (ibid.). With apologies, the material is simply too complex and lengthy to repeat in the space provided here. Chesthetica (presently in version 10.63) is able to evaluate aesthetics in chess move sequences including mates in 3, mates in 4, mates in 5 and 'studies' which are essentially longer sequences that do not necessarily end in mate but are fairly decisive regardless. Full games, however, cannot be evaluated aesthetically because a full chess game (which includes well-documented opening lines and many known endings) is never considered beautiful as a whole. A game is generally considered beautiful due to a particular sequence in it rather than all the moves [1][19].

2.2. The Games Used

For this research, we used two books as sources for the games between men and women. The first book was 'Play Like a Girl: Tactics by 9 Queens' by Jennifer Shahade [20]. She is a two-time U.S. Women's Champion and a Woman Grandmaster (WGM). Given the scarcity of expert-selected games by women, this book was the best resource of its kind that we could find. The second book was the comparable 'The Ultimate Chess Puzzle Book' by John Emms [21]. He is an International Grandmaster and tied winner of the 1997 British Championship. The book features similar noteworthy sequences taken from games primarily between men (though likely not intentionally). All the games selected from these books were processed but many had to be filtered out based on unbiased criteria such as the formats that Chesthetica could process (e.g. one side, typically White, clearly winning by the end of the sequence).

Included therefore were the games featured (by the expert authors) with sequences ending in mate and those where one side had a decisive advantage (i.e. 'study'-like). This could mean, for example, White being ahead by a piece after the sequence, which is a clearly winning advantage, especially between strong players. Note that these sequences were not necessarily forced wins (according to perhaps chess engine analysis done later on the position) but taken as they occurred in the actual game, i.e. as presented in the books as the 'main line' the sequence. The games were entered manually into computer following the standard PGN (portable game notation) format that is supported by most chess programs, including Chesthetica. The Elo ratings of the players at the times of those games, if they were available, were also determined using the 'ChessBase Online Database' (approximately 8 million games at the time of writing).[1]

---

[1] https://database.chessbase.com/?lang=en

The recorded game information would then be fed into Chesthetica, which calculated their aesthetic evaluations. A total of 39 games between women, 40 games between women and men (where the women always won), and 115 games between men were able to be derived from the books and were successfully processed based on our best efforts. It was indeed expected that the games between women would be considerably fewer (given the lack of female participation in chess and the sizes of the sourced books) but in the interest of using all available data, the viable games sequences between men were not reduced arbitrarily. The PGN files for all the games are available online.[2][3][4] Noteworthy is also that, where the Black player won, the sequence was simply 'mirrored' in the PGN as winning with the white pieces instead so Chesthetica could process it. Given that most chess problems are presented with White to play and win, the need to support programmatically also Black winning was deemed unnecessary since its inception. This, however, does not at all affect the aesthetics score calculation. It is only the initial position and the winning sequence – regardless of the colour of the pieces, the names of the players or even their ratings – that matter in obtaining the score. Fig. 1 shows two examples of sequences along with their aesthetic scores.

Marta Litinskaya vs. Anna Zatonskih,
ROM-chT, Sovata
1998
1. Rxc4 Rxc4 2. Qb3 Qc7 3. Rc1 Rc8
4. Nd6 Qxd6 5. Bxc8 1-0
Aesthetics Score: 2.543

Branko Kutuzovic vs. Nenad Levar,
Bizovac Metalis op 07th,
2000
1. Qg7+ Bxg7 2. fxg7+ Kg8 3. Bxh7+ Kxh7
4. gxf8=Q Qe4+ 5. Qf3 Qc2+ 6. Qf2 Qe4+ 7. Kg1 1-0
Aesthetics Score: 2.476

Fig. 1: Examples of the game move sequences and their aesthetic scores.

2.3. Statistical Processing & Results

For statistical purposes, an F-test was used on the Elo ratings and aesthetic scores to determine equality of variance. This then determined whether a two-tailed T-test assuming equal (TTEV) or unequal (TTUV) variances (at the 5% level) was applied to compare the means (and determine statistical significance). For each PGN (i.e. women vs. women, women vs. men and men vs. men) it was first determined whether the Elo ratings of the players were comparable. This means that given the available Elo ratings for the White players and Black players (for each PGN), the ratings, on average, should not be different. For all the groups, i.e. women vs. women (71.8% of ratings for White and 74.4% of ratings for Black could be identified), women vs. men (92.5% of ratings for White and 87.5% of ratings for Black could be identified) and men vs. men (65.2% of ratings for White and 63.5% of ratings for Black could be identified), there was no statistically significant difference.

Specifically, comparing the mean ratings of the female players against the male players (excluding the mixed group), there was no statistically significant difference based on the TTEV test; t(225) = -1.733, P = 0.084. The women averaged an Elo rating of 2,394 and the men averaged an Elo rating of 2,427. In the mixed group (women vs. men), there was also no statistically significant difference between the White and Black player Elo ratings based on the TTEV test; t(70) = -0.528, P = 0.6. The White players averaged an Elo rating of 2,456 and the Black players averaged an Elo rating of 2,478. This meant that players from both

---

[2] http://metalab.uniten.edu.my/~azlan/Other/WCSE2018/men_vs_men.pgn

[3] http://metalab.uniten.edu.my/~azlan/Other/WCSE2018/women_vs_women.pgn

[4] http://metalab.uniten.edu.my/~azlan/Other/WCSE2018/women_vs_men.pgn

genders were equally strong and ruled out playing strength as a factor. Table 1 shows the aesthetic score comparisons.

Table. 1: Mean aesthetic scores of the groups.

Women vs. Women (Group 1)	Women vs. Men (Group 2)	Men vs. Men (Group 3)
1.598	1.821	1.661

There was no statistically significant difference between groups 1 and 3 based on the TTEV test; $t(152) = -0.612$, $P = 0.542$. A single factor analysis of variance (ANOVA) test was performed across all the three groups with no statistically significant difference detected; $F(2, 191) = 1.695$, $p = 0.186$. Finally, groups 1 and 2 were compared with no statistically significant difference detected based on the TTEV test; $t(77) = -1.728$, $P = 0.088$. So aesthetically, the winning sequences taken from games between women and men were no different in quality.

3. Discussion & Conclusions

The results in section 2.3 should be considered more reliable than earlier work [10] not only because playing strength was ruled out as a factor but also because sequences longer than just three-movers were analyzed aesthetically, i.e. the present work included longer mates and study-like endings as well. This research therefore contradicts and supersedes the earlier work. The fact that no aesthetic differences could be detected comparing games by men and women suggests that within the domain of chess, winning is either objectively beautiful or that both men and women have playing approaches that are not unlike each other. In other words, they perceive things similarly. When it comes to the human body, for instance, aesthetic perception between men and women have been shown to vary somewhat [22]. Research into gender perception differences with regard to games and specifically board games like chess is scarcer and the present work hopefully contributes to the literature in that regard.

The implications of this work are that measures used to promote chess to young girls and boys (and to train them) should not be assumed to be different. The same principles of beauty in the game that appeal to males should also appeal to females. If there is any difference it may, in fact, be due to initial differences in playing strength. There would be little doubt, for instance, that a significantly higher rated female player is more likely to win a game against a lower rated male player. It follows therefore that equally or comparably rated male and female players are just as likely or capable of playing and winning in equally impressive or aesthetic ways. The quality of their games, as a whole, should also be comparable. Even so, there still remains a significantly smaller number of brilliancies and award-winning chess compositions (if any) by females. This is likely attributable to the fact that women are underrepresented in the sport, and especially in the domain of chess problem composition. It may also be that females, in general, simply do not enjoy chess as much as males. It is, after all, a war-like game and women are naturally less aggressive and competitive from having lower levels of testosterone; so this may indeed contribute to that discrepancy.

4. Acknowledgements

5. References

[1]   I. Damsky, Chess Brilliancy, Everyman Publishers, London, England, ISBN 978-1857442748, 2002.

[2]   A. Avni, Creative Chess, Everyman Publishers, London, England, ISBN 978-1-85744-149-9, 1998.

[3]   J. Levitt and D. Friedgood, Secrets of Spectacular Chess, 2nd Edition (Expanded), 1 May, Everyman Chess, London, England, ISBN 978-1-85744-551-0, 2008.

[4]   K. Howard, The Enjoyment of Chess Problems, Dover Publications, New York, ISBN 978-0486207421, 1967.

[5]   M. Lipton, R. C. O. Matthews and J. Rice, Chess Problems: Introduction to an Art, Citadel Press, New York, 1965.

[6]   I. Sukhin, Chess Gems: 1000 Combinations You Should Know, Mongoose Press, Massachusetts, USA, ISBN 978-0-9791482-5-5, 2007.

[7]   P. N. Humble, The Aesthetics of Chess: A Reply to Ravilious, British Journal of Aesthetics, vol. 35, no. 4, pp. 390-94, 1995.

[8]   A. Martin, A Male Dominated Game? ChessBase News, Hamburg, Germany. https://en.chessbase.com/post/a-male-dominated-game, 5th February 2018.

[9]   R. McLellan, Chess Educators Explore Connections Between Chess and the Arts, ChessBase News, Hamburg, Germany, https://en.chessbase.com/post/chess-educators-explore-connections-between-chess-and-the-arts, 30th January 2017.

[10] A. Iqbal, Which Gender Plays More Beautiful Chess? *Proceedings of the 4th International Congress on Interdisciplinary Behavior and Social Science 2015* (ICIBSoS 2015), Kazan Federal University, Russia, 22-23 October & Arya Duta Hotel, Jakarta, Indonesia, 7-8 November, pp. 423-427, CRC Press, Taylor & Francis Group, ISBN 9781138028388, CAT# K26611, 2015.

[11] C. Gerdes and P. Gränsmark, Strategic Behavior Across Gender: A Comparison of Female and Male Expert Chess Players, *Labour Economics*, 17(5), pp. 766-775, 2010.

[12] C. F. Chabris and M. E. Glickman, Sex Differences in Intellectual Performance Analysis of a Large Cohort of Competitive Chess Players, *Psychological Science*, 17(12), pp. 1040-1046, 2006.

[13] A. Maass, C. D'Ettole and M. Cadinu, Checkmate? The Role of Gender Stereotypes in the Ultimate Intellectual Sport, European *Journal of Social Psychology*, 38(2), pp. 231-245, 2008.

[14] M. Bilalić, K. Smallbone, P. McLeod and F. Gobet, Why are (the Best) Women So Good at Chess? Participation Rates and Gender Differences in Intellectual Domains, *Proceedings of the Royal Society B: Biological Sciences*, 276(1659), pp. 1161-1165, 2009.

[15] F. Friedel, Chess Gender Debate in the International Press, ChessBase News, Hamburg, Germany, http://en.chessbase.com/post/chess-gender-debate-in-the-international-press, 21st April 2015.

[16] T. Stafford, Chess Players Outperform Expectations When Playing Men, *Psychological Science*, Vol. 29, Issue 3, pp. 429-436, 2018.

[17] N. Vaci and M. Bilalić, Chess Databases as a Research Vehicle in Psychology: Modeling Large Data, *Behavior* Research *Methods*, August, Volume 29, Issue 4, pp. 1227-1240, 2017.

[18] A. Iqbal, H. van der Heijden, M. Guid and A. Makhmali, Evaluating the Aesthetics of Endgame Studies: A Computational Model of Human Aesthetic Perception, *IEEE Transactions on Computational Intelligence and AI in Games*, 4(3), pp. 178-191, 2012.

[19] D. Shenk, The Immortal Game: Or How 32 Carved Pieces On a Board Illuminated Our Understanding of War, Art, Science, and the Human Brain, Anchor Canada, ISBN 978-0385662277, 2007.

[20] J. Shahade, Play Like a Girl: Tactics by 9 Queens, Mongoose Press, ISBN 978-1936277032, 2011.

[21] J. Emms, The Ultimate Chess Puzzle Book, Gambit Publications, ISBN 978-1901983340, 2001.

[22] S. Silveira, K. M. Elvers, K. Fehse and M. Paolini, Gender Differences in the Perception of Personalized Half-Nude Female Bodies, *Frontiers in Psychology*, 8, p. 1529, 2017.