An Empirical Study of Artificial Participants: A Factor Affecting Game Length in Chess

William Bart, University of Minnesota, USA Jacob Ritter, University of Minnesota, USA Nathan Ritter, University of Minnesota, USA

ABSTRACT

This study is an investigation of artificial participants operating in their appropriate environment. The artificial participants in this study are artificial chess players, and their appropriate environment is a chess game. This study is an empirical investigation testing the hypothesis that the length of a chess game is inversely related to the difference in the chess skill levels of the artificial chess players. Five series of 18 chess games were instituted between five pairings of web-based chess engines. The chess engine, chess.com Level 10, was a player in each series and won all 18 games in each series. The opposing players came from the Play Magnus app at five different levels of chess skill. This study provided an investigation of 90 chess games involving artificial chess players. The hypothesis for the study was confirmed. Game length was significantly inversely related to the disparity in chess proficiency between artificial chess players. This is one of the first scientific studies of artificial participants.

KEYWORDS

Advanced Chess Technology, Artificial Intelligence, Complex Games, Game Length, Play Magnus, Quantitative Study, Scientific Research, Stockfish

INTRODUCTION

With new electronic technologies such as those dealing with mobile communications and Robotics comes the emergence of artificial participants that are software-directed entities that can engage in tasks that are typically done by humans and that can serve as participants in empirical research. In certain forms of psychological research, one could thus have human participants as well as artificial participants. With the proliferation of artificial participants, one can inquire how they perform in their appropriate environments and what factors affect that performance.

There are many forms of artificial participants such as robotic waiters programmed to wait on customers in restaurants, social service robots programmed to provide information to customers in shopping malls, security robots programmed to detect and prevent crime in shopping malls, and chess engines programmed to play chess at a high level. All of those entities can serve as participants in empirical research.

DOI: 10.4018/JTA.20210101.oa2

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

One form of artificial participant that has appeal in the restaurant business is the robotic waiter that is intended to serve food to customers in a restaurant. Zhang (2018) reported that robotic waiters have become popular in a host of restaurants in China. The robotic waiter takes the form of a transparent container the size of a microwave oven that is placed on a cart that rolls around the restaurant. When the robotic waiter arrives at a table, it raises its glass lid to reveal the ordered food and then tells the customer to enjoy the meal.

Mishra, Goyal, and Sharma, (2018) reported that robotic waiters are already being used in countries such as China and Japan and that the restaurant businesses in other countries such as India are developing interests in them. However, further improvements in that technology are warranted. For example, there is the issue of navigation in room service or serving at a table. Robotic waiters at present typically follow a track in a restaurant. This approach is termed the line follower technique. One positive aspect to the line follower technique is that robotic waiters following this technique can move about smoothly in a restaurant in a balanced manner. One problem is that the robotic waiter following the line follower technique typically needs to be stopped by a customer who then takes the food from the robotic waiter and places the food onto a table being used by the customer. Robotic waiters are not advanced enough to serve food on a table as well as a human.

In a subsequent publication on the same topic, Mishra, Goyal, Sharma, and Gupta (2021) indicated that the task of a robotic waiter is to identify and stop at the correct table of a restaurant customer with the food order. But robotic waiters do not always behave the way intended. They at times stop at some distance from the desired table or even stop at the wrong table which indicates a detection problem in the robotic waiter technology.

A possible solution to the detection problem is the use of RFID technology (Ahuja and Potti, 2010; Finkenzeller, 2010), which refers to Radio Frequency Identification Technology. RFID is a technology used to scan products as in grocery stores. It is superior to reading barcodes in that it permits the scanning of objects at a distance. Robotic waiters in restaurants equipped with RFID tags and RFID tag readers allow accuracy in the detection by a robotic waiter of the correct table in the serving of food. But the problem of how to place the food on the customer table correctly remains.

Although some of the behaviors of robotic waiters have improved, robotic behaviors still do not behave completely as intended. More research is required to address the problem of correct food placement. Robotic waiters are examples of artificial participants that behave in specific ways within the environment of restaurant. They have been programmed to do their job fairly well, but further improvements are warranted.

One shortcoming to the research on robotic waiters is the lack of use of the scientific method in the research. Rarely is a hypothesis cited and then tested empirically with the collection of relevant data and the subsequent analysis of the data with the use of appropriate statistical methods. Research on robotic methods would likely benefit from the use of rigorous scientific analysis.

Another category of artificial participants is that of robots that operate in shopping malls that can provide guidance to shoppers or that surveil the shopping mall. Niemelä, Heikkilä, Lammi, and Oksman (2017) reported on social service robots that provide information to customers in shopping malls. They interviewed eight retailers and three shopping mall managers to assess their responses to the robots. They found that social service robots that can engage with customers on an emotional level will likely contribute to business activity in the shopping mall. They also reported that consumers tend to respond favorably to these robots.

In a subsequent publication on social service robots, Niemelä, Heikkilä, Lammi, and Oksman (2019) studied the reactions of customers, store managers, and mall supervisors to shopping mall robots. The robot studied was called Pepper. In general, the people, who were questioned, presented positive views of the robot. The robot was valued to the extent to which it was useful and periodically entertaining. A successful social service robot should have a well-developed dialog capacity to provide information to customers and to work with personnel in a collaborative manner. The robot

was also valued to the extent to which it was viewed as an agent to increase store sales and to decrease shopping mall costs.

The other category of shopping mall robots is that of security robots that provide surveillance of the shopping mall. Trovato, Lopez, Paredes, and Cuellar (2017) reported on an experiment with 55 human participants assessing their views of a security robot in a shopping mall. They found that the robot was perceived to be effective to the extent to which its behavior was perceived to reflect authority and kindness. They contended that their results inform decisions regarding the possible purchase and usage of security robots in real situations in shopping malls.

Social service robots and security robots are examples of artificial participants that behave in specific ways within the environment of shopping malls and similar locations. They have been programmed to do certain tasks, but it is not certain how well they do those tasks. Social service robots are effective only if their dialog capacities are very well developed. Security robots are effective only if their capacities to perceive problematic conditions are very well developed. Both types of robots will likely benefit from further improvement to address their major needs – i.e., improved dialog capacities in the case of social service robots and improved problematic condition detection capacity in the case of security robots.

The research on social service robots and security robots used in shopping malls is beset with the same shortcoming as beset the research on robotic waiters, namely a lack of studies that make use of the scientific method in research. The use and subsequent analysis of interview data as in the study by Niemelä, Heikkilä, Lammi, and Oksman (2019) provides a description of the reactions of customers of shopping malls, business managers, and shopping mall supervisors in one limited case, but the findings of an essentially descriptive study may not be generalizable to other groups of customers, business managers, and supervisors involved with other shopping malls. Only with the use of the scientific method in which a hypothesis is cited and then tested empirically with the collection of relevant data and the subsequent analysis of the data with the use of appropriate statistical methods can the study of shopping mall robots produce generalizable results. Research on shopping mall robots would clearly benefit from the use of rigorous scientific analysis.

Another domain of artificial participants that has available an indeterminate number of artificial participants is the game of chess which has been the setting for many interesting studies. For example, Levene and Bar-Ilhan (2007) investigated the similarity of opening moves made by chess engines (i.e., artificial chess players) in chess engine vs chess engine chess competitions with opening moves made by humans in human vs human chess competitions. They analyzed opening moves resulting from 26 chessboard positions and, using several nonparametric statistics, found a strong similarity of opening moves chosen by humans with those chosen by the chess engines.

However, the human participants in the chess competitions that Levene and Bar-Ilhan used in their study were mostly excellent with many being professional chess players and the artificial participants in the chess engine competitions played chess at high levels of chess proficiency. It is not clear if the similarity in opening move selection would be maintained in comparisons of typical opening moves made by less competent human chess players and less competent artificial chess players. The similarity that they found may be positively related to the degree of similarity between the levels of chess competency of the human chess players and the levels of chess competency of the artificial chess players in their respective chess competitions. Scientific research of that conjecture would determine how much scientific merit that conjecture has.

Guld and Bratko (2011) used four different chess engines to analyze large sets of human chess move decisions and determined highly similar rankings of the chess performances of the chess players with any of the programs. They concluded that all of the four programs agreed as to which human player was stronger than another human player and that the chess engines that they studied such as CRAFTY could be used to estimate the level of chess proficiency of chess players. One issue with their study is that the chess players whom they studied were outstanding chess players and that the chess engines were able to play at high levels of chess proficiency. It is unclear whether their findings would hold for less competent human chess players and less proficient chess engines.

Ferreira (2013) did an interesting study to investigate the relationship of depth of search is related to the level of chess skill of a human chess player. He used the powerful chess program HOUDINI at different levels of search depth to analyze the quality of play in games played at the Candidates Tournament of 2013 to determine the next Challenger to play the World Chess Champion. Depth of search is defined in terms of the number of plies used to determine a move. For example, if a chess player looks at a chess position and then considers an array of moves and their resulting positions before deciding on a move, then the chess player employed a search with a depth of one ply. If the chess player looks at a chess position and then considers the moves that the player could make and then the resulting moves that the opponent could make along with the resulting positions before deciding on a move, then the chess player employed a search with a depth of two plies. Ferreira found that HOUDINI operating at a depth of search of 20 plies played chess similar to the chess play of the outstanding huma chess players in the Candidates Tournament. He concluded that HOUDINI operating at a depth of search of 20 plies highly correlated with the Elo ratings (international measure of chess skill) of the human chess players. That result clearly suggests that high-level human chess players can think deeply when deciding on chess moves.

McIlroy-Young, Sen, Kleinberg, and Anderson (2020) reported on a very exciting project. They determined that AlphaZero, a very powerful chess engine that uses artificial intelligence, does not predict chess moves made by human chess players. To better mirror human chess moves, they developed and introduced Maia, a chess engine using artificial intelligence, that predicts human chess moves. In other words, Maia makes chess moves similar to those of a human chess player after processing a set of the chess moves made by the human chess player with the use of a neural network. This is exciting news, because Maia and similar programs will permit scientists to better understand with precision how human and artificial chess players play chess at different levels of chess skill.

Although there are many artificial participants available in chess including the many chess programs (also called chess software or chess engines) such as Stockfish, Komodo 12, and Fritz 17 that can be set at different levels of chess competency, the focus of chess research involving artificial participants has tended to be on artificial chess players that operate at very high levels of chess skill.

Arguably the most famous artificial participant in chess is Deep Blue. On February 10, 1996, the IBM chess computer Deep Blue defeated Gary Kasparov and thus became the first chess computer to defeat a reigning World Chess Champion in a game of chess under tournament conditions. Although IBM has not made public the code for the Deep Blue software, the domain of chess software has flourished. Not only are there now chess programs stronger than Deep Blue, but there are now weakened versions of strong chess programs that are at different levels of chess competency. For example, the chess website lichess.org allows one to play the smart chess software Stockfish 12 at eight levels of chess proficiency with Stockfish 12 level 1 being the weakest artificial chess player available at the lichess.org website and Stockfish 12 level 8 being the strongest artificial chess player available at the same website.

The technological advancements in chess software have been amazing. Not only is there now smart chess software that provides artificial chess players that are superior at chess to any human players, but there is now a plethora of artificial chess players at different levels of chess competency that are available for psychological study that could shed light on how artificial participants behave. In fact, there is now an indeterminate number of artificial chess players available to serve as participants in psychological research.

One way that chess software technology has advanced is through the introduction of artificial intelligence (AI) concepts in the design of chess software. AlphaZero is the first chess software that incorporated artificial intelligence concepts (Silver et al., 2018). Although it has a different design than other forms of chess software, it has demonstrated superiority at playing chess when playing very strong artificial chess players.

As recognition of the sophisticated technological advancement of AlphaZero, let us consider the case of the winner of the 2019 Averbakh-Boleslavsky Award for the best chess book for 2019. FIDE (Fédération Internationale des Échecs) is the International Chess Federation located in Paris that oversees all national chess organizations and international chess competitions. Annually FIDE provides a prestigious award termed the Averbakh-Boleslavsky Award to the best chess book of the year. The recipient of that award *Game Changer* written by Matthew Sadler and Natasha Regan (Sadler & Regan, 2019) won the prestigious Averbakh-Boleslavsky Award for 2019. Chess Grandmaster Artur Jusupov, a member of the award jury, reported that the book is a unique project that combines human achievements with remarkable development in AI and opens a new approach to the beloved game.

In that book, Sandler and Regan reported on a series of games between two formidable chess engines (i.e., chess-playing programs), Stockfish 8, a leading open-source chess engine, and AlphaZero. Stockfish 8 had much chess knowledge including knowledge of openings, middle game tactics, and endgame solutions, programmed into it by skilled programmers.

AlphaZero, on the other hand, was only provided the basic rules of chess and, after playing itself frequently for about four hours, attained tremendous chess skill. Both chess engines have skills at playing chess that eclipse those of the vast majority of expert human chess players. AlphaZero was deemed superior to Stockfish 8 having defeated Stockfish 8 in a majority of their games played.

As an indication of the levels of chess prowess of both Alpha Zero and Stockfish 8, let us consider their estimated Elo ratings. The Elo rating is the international metric used to determine the level of chess proficiency of a chess player. The Elo rating is the invention of Arpad Elo, a physicist and avid chess player. An Elo rating of 0 would indicate the absence of any chess skill. The Elo rating of Magnus Carlsen, present World Chess Champion, as of March, 2021, was 2847 https://en.chessbase. com/post/fide-ratings-march-2021. Silver, et. al. (2018) estimated the Elo rating of AlphaZero to be 3200 and the Elo rating of Stockfish 5 to be 3100.

One noteworthy feature of the series of games between AlphaZero and Stockfish 8 is that the length of the games in terms of the number of moves played in a game tended to be large. For example, one game, Stockfish 8 vs AlphaZero, London 8, 2018, lasted 142 moves (Sadler & Regan, 2018)! Many of the games between the two powerful chess engines and discussed by Sadler and Regan (2018) lasted more than 50 moves. One explanation for these long game lengths may be that the two artificial opponents were similar in chess skill.

This leads to the following research question investigated in this study: What is the relationship between the disparity between the chess skill levels of two artificial chess players and the length of their games in moves? In other words, if there is an increase in the disparity between the chess skill levels of two artificial chess players, will the length of their chess games in terms of the number of moves to a final outcome (win, draw, or loss) decrease and conversely if there a decrease in the disparity between the chess skill levels of two artificial chess players, will the length of their chess games in terms of the number of moves to a final outcome increase?

That research question engenders the following hypothesis: the length of a chess game in terms of the number of moves to a final outcome is inversely related to the disparity in chess skill between the two artificial chess players – i.e., the length of a chess game decreases as the difference in the chess skill levels of the artificial chess players increases.

To test scientifically that hypothesis, one cannot readily use AlphaZero, because that chess engine is not readily available. In addition, one would need chess engines that vary in chess competency to provide a range in disparity in player competency levels. Both AlphaZero and Stockfish 8 are simply too strong and too similar in chess competency levels to permit a careful examination of those questions.

One needs pairings of artificial chess players that vary in how they differ in their levels of chess competency. Chess engines available at the chess.com website and the Play Magnus app provide artificial chess players that vary in levels of chess competency and that could produce pairings of artificial chess players engaging in chess games in which there is variation in disparity in player chess skill. As a result, there will be some games in which the artificial opponents will have similar levels

of chess competency and other games in which the artificial chess opponents will have markedly dissimilar levels of chess competency.

METHOD

Sample

There were six participants in this study: (1) chess.com Level 10; (2) Play Magnus Age 10; (3) Play Magnus Age 9; (4) Play Magnus Age 8; (5) Play Magnus Age 7; and (6) Play Magnus Age 6. The artificial participants used in this study do not make use of artificial intelligence technology that Maia uses. chess.com Level 10 is a chess engine available at the chess.com website, https://www.chess.com/play/computer. Play Magnus Age 10-6 are chess engines available at the Play Magnus website on Google Play, https://play.google.com/store/apps/details?id=pl.mw.playmagnus&hl=en_US.

An estimate of the Elo rating for chess.com Level 10 is 2300 (https://www.chess.com/forum/ view/general/approximate-ratings-of-chess-com-computer-levels). Estimates of the Elo ratings of the Play Magnus players are the following: (1) Play Magnus Age 10 – 1700 Elo rating; (2) Play Magnus Age 9 – 1500 Elo rating; (3) Play Magnus Age 8 – 1200 Elo rating; (4) Play Magnus Age 7 – 1000 Elo rating; and (5) Play Magnus Age 6 – 800 Elo rating (https://www.reddit.com/r/chess/ comments/2qcv95/what_is_the_strength_of_the_play_magnus_app_at/). It is important to note that these Elo rating estimates are only estimates listed on the Internet by chess enthusiasts and have not been validated through a series of games with opponents with validated Elo ratings.

Measures

The primary measure used in this study is the length of each game played in terms of number of moves to a final outcome. There were five series of games with each series consisting of 18 games; thus, each series of games resulted in 18 values of game length. The total number of game lengths in this study is thus 90.

Each series consisted of 18 games based on the following rationale. There are three outcomes in a chess game: win, draw, and loss. Assuming that each outcome has an equal probability (i.e., .33), the probability (p) of a win = 1/3. If the number of games (n) =18, the expected number of wins = 6 (np = 18 x 1/3 = 6) with a standard deviation of 2 (sqrt(npq) = sqrt(18 x 1/3 x 2/3) = sqrt(4) = 2). The 95% confidence interval in this case would be $np \pm 1.96 x$ sqrt(npq) = $6 \pm 1.96 x 2 = 6 \pm 3.92 = (2.08, 9.92)$. 18 wins in one series of 18 games would clearly be beyond a 95% confidence interval assuming the probability of a win = 1/3.

If one assumes the probability of a win = .50, then the 99% confidence interval would be np $\pm 2.58 \text{ x sqrt}(npq) = 18 \text{ x } .50 \pm 2.58 \text{ x sqrt}(18 \text{ x } .5 \text{ x } .5) = 9 \pm 2.58 \text{ x sqrt}(4.5) = 9 \pm 2.58 \text{ x } 2.12$ = 9 $\pm 5.47 = (3.53, 14.47)$. 18 wins in one series of 18 games would clearly be beyond even a 99% confidence interval assuming the probability of a win = .50. If one player defeated an opponent 18 games, then that would provide substantial evidence that the one player is clearly superior to the other opponent in chess skill.

Design

The five series of 18 games each had the following pairings: (1) Series1 with chess.com Level 10 vs Play Magnus Age 10; (2) Series2 with chess.com Level 10 vs Play Magnus Age 9; (3) Series3 with chess.com Level 10 vs Play Magnus Age 8; (4) Series4 with chess.com Level 10 vs Play Magnus Age 7; and (5) Series5 with chess.com Level 10 vs Play Magnus Age 6. The participants thus engaged in five series of games.

In each series, the two artificial opponents played 18 games against each other alternating colors for successive games. As a result, each opponent played nine games as White and nine games as Black. All moves were made within seconds.

Series	Players	М	SD
1	chess.com Level 10 vs Play Magnus Age 10	46.61	10.39
2	chess.com Level 10 vs Play Magnus Age 9	37.50	9.48
3	chess.com Level 10 vs Play Magnus Age 8	31.22	5.17
4	chess.com Level 10 vs Play Magnus Age 7	26.72	4.80
5	chess.com Level 10 vs Play Magnus Age 6	19.94	4.95

Table 1. Means and Standard Deviations of Game Lengths for Five Series of Games

For each game, the "Play Magnus" app on a phone and the computer on <u>chess.com</u> with a laptop were used. The following procedure was followed to play the games. After navigating to <u>chess.com</u>, click on "Play" on the left sidebar and then click "Computer" on the drop-down menu that appears. Then click on the "+" sign to change the level of difficulty to level 10. After opening the "Play Magnus" app on a phone, select "Play Me" and then "Choose My Age". Then click on the left arrows or right arrows to decrease or increase the age (or swipe left or right to change the age).

For every set of data collected, Play Magnus played as White during the odd-numbered games and as Black during the even-numbered games. Thus, for the odd-numbered games, Play Magnus would make the first move on the "Play Magnus" app on the phone. Then make the Play Magnus move on the laptop. Next, the <u>chess.com</u> computer would make its first move as Black on the laptop and then make that move on the "Play Magnus" app on the phone. Repeat this process for each move until the game had finished.

For the even-numbered games, the <u>chess.com</u> computer would make its first move as White on the laptop. Next, make the computer's move on the "Play Magnus" app on the phone. Play Magnus would then make its first move as Black on the "Play Magnus" app on the phone, and then make the Play Magnus move on the laptop. Repeat this process until the game had concluded and record the number of moves in the game (i.e., game length).

RESULTS

Chess.com Level 10 won all 18 games in each series, demonstrating its superiority in chess proficiency to the other five artificial opponents, Play Magnus Age 10-6. Table 1 presents descriptive statistics for game lengths in the five series of 18 games.

The 18 games in Series1 that had chess.com Level 10 and Play Magnus Age 10 as players (M = 46.61, SD = 10.39) had, on the average, significantly longer game lengths, t(34) = 2.75, p = .005, than the 18 games in Series2 that had chess.com Level 10 and Play Magnus 9 as players (M = 37.50, SD = 9.48). The effect size was considered to be very large (Cohen, 1988), with Cohen's d = 0.916. The two-tailed *t*-test post-hoc statistical power for this result is 0.759.

The 18 games in Series2 that had chess.com Level 10 and Play Magnus Age 9 as players (M = 37.50, SD = 9.48) had, on the average, significantly longer game lengths, t(34) = 2.47, p = .01, than the 18 games in Series3 that had chess.com Level 10 and Play Magnus 8 as players (M = 31.22, SD = 5.17). The effect size was large, with Cohen's d = 0.822 and Glass's delta = 1.215; Glass's delta is preferable as a measure of effect size, because the two standard deviations are significantly different. The *F*-test for the significance of the difference between the variances of the game lengths for the

Journal of Technological Advancements

Volume 1 • Issue 1 • January-December 2021

Figure 1. Box plots of five series of chess



two series of games indicates that the two standard deviations are significantly different, F(17,17) = 3.36, p < .008. The two-tailed *t*-test post-hoc statistical power for this result is 0.666.

The 18 games in Series3 that had chess.com Level 10 and Play Magnus Age 8 as players (M = 31.22, SD = 5.17) had, on the average, significantly longer game lengths, t(34) = 2.71, p = .01, than the 18 games in Series4 that had chess.com Level 10 and Play Magnus 7 as players (M = 26.72, SD = 4.80). The effect size was again very large, with Cohen's d = 0.902. The two-tailed t-test post-hoc statistical power for this result is 0.746.

The 18 games in Series4 that had chess.com Level 10 and Play Magnus Age 7 as players (M = 26.72, SD = 4.80) had, on the average, significantly longer game lengths, t(34) = 4.17, p = .0002, than the 18 games in Series5 that had chess.com Level 10 and Play Magnus 6 as players (M = 19.94, SD = 4.95) The effect size was very large, with Cohen's d = 1.391. The two-tailed *t*-test post-hoc statistical power for this result is 0.981.

These results provide clear statistical evidence that confirms the hypothesis that the length of a chess game in terms of the number of moves to a final outcome is inversely related to the disparity in chess skill between the two artificial chess players – i.e., the length of a chess game decreases as the difference in the chess skill levels of the artificial chess players increases.

As Figure 1 pictorially indicates, game length tends to decrease as the age level of the Play Magnus artificial player decreases. To further investigate that relation, a. simple linear regression was calculated to predict game length, a dependent variable, based on the age posited for the Play Magnus artificial player, the independent variable. A significant equation was found (F(1,88) = 137.23, p < .00), with an R2 = .61. Participants' predicted game length in moves = (19.59 + 6.51x (age level of the Play Magnus artificial player)) moves, when age level of the Play Magnus artificial player is measured in years of age posited for the Play Magnus artificial player.

Game length increased 6.51 moves for each year of age of the Play Magnus opponent. The age level of the Play Magnus artificial player accounts for 61% of the variance among the game lengths for games between chess.com Level 10 and Play Magnus Age 10-6. The age level of the Play Magnus artificial player thus has a major influence on game length for games involving that player.

CONCLUSION

Summary

Sadler and Regan (2019) provided an extensive qualitative analysis of the games between AlphaZero and Stockfish 8, two formidable chess engines. They discussed the strategies and tactics used by

both artificial opponents. Their book was not a quantitative study involving statistical analysis of measurable features of games involving those two artificial chess titans.

This present study is a quantitative study of measurable features of games involving weaker artificial chess players. In fact, it is likely one of the first quantitative scientific studies, if not the first quantitative scientific study, of the games of artificial players.

Certain findings emanate from this study. One, the artificial chess player, chess.com Level 10, dominated all of its artificial opponents by winning all 90 games. There was not even one draw. This clearly indicated that an artificial chess player can be clearly better than an array of other competent artificial chess players. Two, the variation in game length amidst a series of games between artificial chess players used in this study tended to decrease as the disparity in chess skill between the artificial chess players increased. Three, game lengths tended to decrease as the disparity in chess skill between artificial opponents used in this study increased.

Future Lines of Research

Certain lines of future inquiry are suggested by this study. One, this study could be replicated to determine how reliable its findings are. Two, this study could be replicated using other chess engines and alternative time controls to determine how generalizable the findings of this study are across other artificial chess players and time controls. Three, similar studies of weaker artificial chess players could be designed and implemented to provide information on the strategies and tactics used by the players as well as information on the strengths and weaknesses of the artificial players. Qualitative analyses of features of games involving weaker artificial chess players could ensue and the findings could be compared to similar studies involving human chess players. Such inquiry could result in the improvement of chess training programs that make use of chess engines to improve chess proficiency levels among humans in schools, homes, and other settings.

This study provides an example of how technological advancements in software production that can do intelligent tasks has led to the creation of multiple artificial participants available for psychological research. With artificial participants, one can devise multiple psychological experiments to investigate features of the artificial participants such as how they behave and interact with each other. This study provides an example of a psychological experiment involving artificial participants. This study is an experimental study investigating and heralding the study of the psychology of artificial participants – with special attention

REFERENCES

Ahuja, S., & Potti, P. (2010). An Introduction to RFID Technology. *Communications and Network*, 2(3), 183–186. doi:10.4236/cn.2010.23026

Cards, Radio Frequency Identification and near-Field Communication. (Third Edition). John Wiley & Sons.

Chess Base. (2021, March 1). Elo rating of Magnus Carlsen.https://en.chessbase.com/post/fide-ratings-march-2021 "

chess.com. (2020). chess.com Level 10. https://www.chess.com/play/computer

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Routledge.

Ferreira, D. R. (2013, June 1). The Impact of the Search Depth on Chess Playing Strength. *International Computer Games Journal*, *36*(2), 67–80. doi:10.3233/ICG-2013-36202

Finkenzeller, K. (2010). RFID Handbook: Fundamentals and Applications in Contactless Smart

Guid, M., & Bratko, I. (2011, June 1). Using Heuristic-Search Based Engines for Estimating Human Skill at Chess. *International Computer Games Journal*, *34*(2), 71–81. doi:10.3233/ICG-2011-34204

Levene, M., & Bar-Ilan, J. (2007, September). Comparing Typical Opening Move Choices Made by Humans and Chess Engines. *The Computer Journal*, 50(5), 567–573. https://doi.org/10.1093/comjnl/bxm025

McIlroy-Young, R., Sen, S., Kleinberg, J., & Anderson, A. (2020). Aligning Superman AI with Human Behavior: Chess as a Model System. (Paper presentation). *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Virtual 2020.* doi: 10.1145/3394486.3403219

Mishra, N., Goyal, D., & Sharma, A. D. (2018, April 20). Issues in Existing Robotic Service in Restaurants and Hotels. *Proceedings of 3rd International* Conference on Internet of Things and Connected Technologies (ICIoTCT) held at Malaviya National Institute of Technology, Jaipur (India) on March 26-27, 2018, 10.2139/ssrn.3166508

Mishra, N., Goyal, D., Sharma, A. D., & Gupta, A. K. (2021). Robotic Waiters: Detecting Table to Serve Using RFID Tags. In D. Goyal, A. K. Gupta, V. Piuri, M. Ganzha, & M. Paprzycki (Eds.), Proceedings of the Second International Conference on Information Management and Machine Intelligence. Lecture Notes in Networks and Systems, 166. Springer., https://doi.org/10.1007/978-981-15-9689-6_63.

Niemelä, M., Heikkilä, P., Lammi, H., & Oksman, V. (2017). Shopping Mall Robots Opportunities and Constraints from the Retailer and Manager Perspective. In A. Kheddar, E. Yoshida, G. Shizhi, K. Suzuki, J. Cabibihan, F. Eyssel, & H. He (Eds.), Lecture Notes in Computer Science: Vol. 10652. Social Robotics. 9th International Conference on Social Robotics 2017. Springer., https://doi.org/10.1007/978-3-319-70022-9_48.

Niemelä, M., Heikkilä, P., Lammi, H., & Oksman, V. (2019). A Social Robot in a Shopping Mall: Studies on Acceptance and Stakeholder Expectations. In O. Korn (Ed.), Social Robots: Technological, Societal and Ethical Aspects of Human-Robot Interaction. Human–Computer Interaction Series. Springer., https://doi.org/10.1007/978-3-030-17107-0_7.

Play Magnus. (2020). *Play Magnus Age 10-6* https://play.google.com/store/apps/details?id=pl. mw.playmagnus&hl=en_US

Reddit.com. (2020). *Estimates of the Elo ratings of Play Magnus players*. https://www.reddit.com/r/chess/ comments/2qcv95/what_is_the_strength_of_the_play_magnus_app_at/

Sadler, M., & Regan, N. (2019). Game changer: AlphaZero's groundbreaking chess strategies and the promise of AI. New in Chess.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, *362*(6419), 1140–114. doi:10.1126/science.aar6404 PMID:30523106

Trovato, G., Lopez, A., Paredes, R., & Cuellar, F. (2017), Security and guidance: Two roles for a humanoid robot in an interaction experiment, *26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Lisbon, Portugal, 230-235, doi:0.1109/ROMAN.2017.8172307

Zhang, A. (2018, August 5). *Chip labour: robots replace waiters in china restaurant*. https://phys.org/news/2018-08-chip-labour-robots-waiters- china.html

William Bart (PhD) is a Professor of Educational Psychology at the University of Minnesota and the Editor-in-Chief of the International Journal of Gaming and Computer-Mediated Simulations (IJGCMS).

Jacob Ritter is an officer of a chess club and a chess enthusiast.

Nathan Ritter is a chess enthusiast and a chess researcher.