Human Learning from Artificial Intelligence: Evidence from Human Go Players' Decisions after AlphaGo

Minkyu Shin (minkyu.shin@yale.edu)
Yale University, New Haven, CT 06511 USA

Jin Kim (jin.m.kim@yale.edu)
Yale University, New Haven, CT 06511 USA

Minkyung Kim (Minkyung_Kim@kenan-flagler.unc.edu)
University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 USA

Abstract

Although Artificial Intelligence (AI) is expected to outperform humans in many domains of decision-making, the process by which AI arrives at its superior decisions is often hidden and too complex for humans to fully grasp. As a result, humans may find it difficult to learn from AI, and accordingly, our knowledge about whether and how humans learn from AI is also limited. In this paper, we aim to expand our understanding by examining human decision-making in the board game Go. Our analysis of 1.3 million move decisions made by professional Go players suggests that people learned to make decisions like AI after they observe *reasoning processes* of AI, rather than mere *actions* of AI. Follow-up analyses compared the decision quality of two groups of players: those who had access to AI programs and those who did not. In line with the initial results, decision quality significantly improved for the players with AI access after they gained access to *reasoning processes* of AI, but not for the players without AI access. Our results demonstrate that humans can learn from AI even in a complex domain where the computation process of AI is also complicated.

Keywords: Human Learning; Artificial Intelligence; Measure of Learning; Decision Quality

Introduction

This paper presents empirical evidence that human experts learn to make decisions like superhuman Artificial Intelligence (AI) in the context of the board game Go[1]. As AlphaGo demonstrated its superhuman performance by defeating a Go world champion in March 2016, AI is expected to outperform humans in many domains of complex decision-making. However, it is not clear whether humans learn from the output of AI programs and benefit from it. Although information provided by AI algorithms could be useful to humans, the black-box nature of AI can lead humans to misinterpret the output and make decisions no better than before. Examining the impact of AlphaGo on human decision-making, we study whether human experts improve their decisions and what features of AI are essential for them to learn from AI. Hereafter, we refer to improvements in human experts' decision quality after the introduction of AI programs as the *human learning from AI*.

We investigate decision-making problems in the game of Go for two reasons. First, it is one of the first domains in which AI achieved superhuman performance in a complex decision-making problem[2]. This superhuman performance is a useful feature of AI when studying the human learning from AI, primarily because it encourages humans to study decisions of AI and secondarily because superhuman AI can evaluate and track the degree to which human decisions are inferior to its own decisions.

Another reason we examine the game of Go is that a game is an effective setting to test how humans interact with AI and adapt their decision-making. The goal of a game is usually well-defined and human players choose various actions to achieve the goal. Those actions, or any decisions[3], and the resulting changes in the environment often get recorded in a database. Using these unique features of a game, researchers have studied various aspects of human decision-making, from error correction to skill acquisition (Biswas, 2015; Regan, Biswas, & Zhou, 2014; Stafford & Dewar, 2014; Strittmatter, Sunde, & Zegners, 2020; Tsividis, Pouncy, Xu, Tenenbaum, & Gershman, 2017).

Our empirical strategy to study the human learning from AI is as follows. First, we devise a simple and intuitive measure to quantify the human learning from AI (i.e., *how much* the quality of human decisions changed after introduction of the AI programs). Next, we use this measure to estimate human learning on a rich data set that includes outcomes of official matches between professional Go players (N>30K) as well as every move decision made by the players in each of the matches (N>1.3 MM). The data spans a period both before and after the AI programs (such as AlphaGo) were introduced. Finally, we use mandatory military service in South Korea as a natural experiment. Because all Korean males are required to serve in the military for 18-24 months, some male players were isolated from the society and did not have access to the AI programs. We compare those who had access to AI programs with those who did not and estimate how much the AI programs contributed to the human learning from AI.

We find that merely observing AI's *actions* (i.e., *decisions*) may not bring a meaningful improvement in human decision-

[1] Go is a board game between two players who take turns placing "stones" of their color (black or white) on a 19×19 grid of lines. The game's objective is to surround a larger territory on the board than the opponent by completely enclosing it with one's stones.

[2] In contrast to other relatively simple games such as checkers, Go presents arguably the most complex task, which explains why AlphaGo defeating a top human expert was seen as a major breakthrough for artificial intelligence.

[3] We use the terms "actions" and "decisions" interchangeably in this paper.

making. Observing AI's *reasoning processes*, however, does seem to improve human decision-making. Our finding is consistent with previous research based on constructivism learning theory (Resnick, 2018).

Background

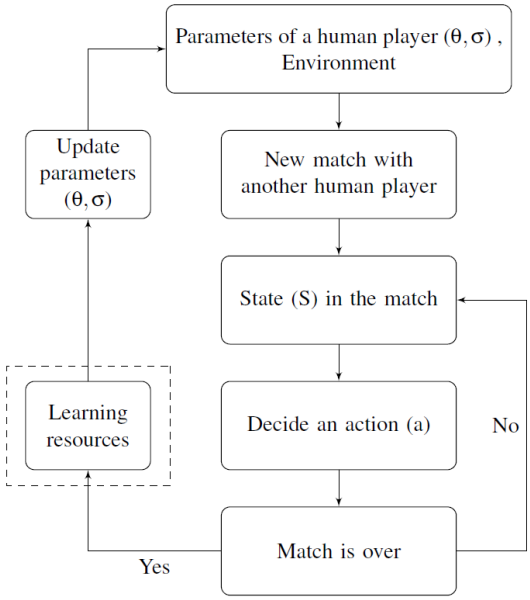
The emergence of AI transformed the way human learning occurred in Go. Before there were AI programs, human players learned Go strategies by reviewing other players' move decisions in tournaments and discussing the strategies with other human players, as illustrated in Figure 1(a)[4]. This way of learning, however, changed after AlphaGo defeated a human Go champion by employing novel and unorthodox tactics. AlphaGo's demonstration of unfamiliar but effective strategies motivated human players to discuss and learn the strategies of AI in addition to those of top human players. Moreover, human players began spending more time playing against AI programs to improve their game, as illustrated in 1(b)[5].

(a) Human Learning Before AI (b) Human Learning After AI

Figure 1: Illustration of change of the way human Go players learn winning strategies after the emergence of AI programs

Two kinds of AI programs were released sequentially, about 1.5 years apart, and we leverage the timing of these events to gain insight about the human learning from AI. In this paper, we focus on AlphaGo's victory over a human Go champion on March 15, 2016 which marks the release of *AI programs that reveal AI actions*. Likewise, we focus on the release of open-source Leela Zero on October 25, 2017 which marks the release of *open-source AI programs that reveal AI reasoning processes*.

These two AI programs feature different types of learning resources. The first AI program, AlphaGo and its subsequent versions, revealed only actions of AI (i.e., sequences of positions where AI placed stones). After observing novel AlphaGo actions, human players discussed among themselves to figure out AlphaGo's intent or subsequent actions that AlphaGo would have taken. However, human players could only make guesses about these. The second AI program, namely open-source AI programs like Leela Zero, revealed not only the actions of AI, but also its reasoning processes. For example, Leela Zero and its ilk showed a set of possible moves that AI considered before making its final move decision. For each of these possible moves, the open-source AI programs calculated the win probability[6] associated with the move and showed how the game will be played following the move (i.e., a series of optimal decisions[7] expected by AI). The online supplementary materials for this paper discuss the information that Leela Zero provided to human players (https://osf.io/skuy5).



Period	Learning resources
Before AI	Human actions
After AlphaGo	Human actions + AI actions
After Opensource AI	Human actions + AI strategy

Figure 2: Simplified Human Learning Process in Go

Before we discuss our method and results, it would be helpful to illustrate how learning typically occurs in Go, and which parts of the learning process were likely affected by the aforementioned AI programs. As shown in Figure 2, learning begins with a human player possessing (1) an evaluation parameter (θ) that assesses a state on the board (i.e., how the pattern of black and white stones placed on the board is advantageous to the player or the opponent) and (2) a strategy parameter (σ) that encompasses a set of decision rules. In a match against another player, the player uses their evaluation parameter to evaluate states (S) and then uses their strategy parameter to take actions (a), i.e., make decisions. After the

[4]"The power of Korean Go? Joint research!" Mar/09/2011 (donga.com/news/Culture/article/all/20110309/35417308/1)

[5]"Hone your Go skills with AI instructors" Feb/26/2019 (munhwa.com/news/view.html?no=2019022601032103009001)

[6]Although we use the term *win probability*, its exact interpretation is more subtle. Still it is used in the Go community so we use this term for convenience.

[7]In this paper, an "optimal decision" refers to the best decision that the agent (human or AI) can make in the given state, rather than the true best decision for the given state. We are interested in learning about how humans make decisions like AI that is far superior to humans, rather than in learning about how humans learn to make *best* decisions in the strict sense of the words.

match ends, the player reviews their own move decisions and those of their opponent, and updates their evaluation parameter (θ; e.g., "The opponent seemed to be in a strong position in State S but they actually had several vulnerable spots") and strategy parameter (σ; e.g., "After some contemplation, I realize that I should have attacked the opponent's territory from the other side").

Before AI programs were available, human players' learning resources consisted only of their own or other human players' move decisions. However, with the advent of the two AI programs, the learning resources available to them changed substantially: AlphaGo introduced actions of superhuman AI to the mix and Leela Zero introduced a chance to study the two decision parameters of AI, σ^{AI} and θ^{AI}. This latter addition to the learning resources by Leela Zero was especially important, because human players could now directly observe how AI evaluates states (e.g., win probability for any states) and how AI sets up a strategy (e.g., sequences of AI decisions in response to the opponent's moves).

Measure

We develop a measure called the *Human-AI Gap* to compare the quality of decisions by humans to the quality of decisions generated by an AI program. Defining the quality of decisions can be a challenge, however, because consequences of decisions are hard to pin down in a high-dimensional state space. Fortunately, modern AI programs based on deep reinforcement learning can not only generate decisions of superhuman quality but also evaluate the quality of any decision. Specifically, AI's value network evaluates states, or situations, in a game, allowing us to evaluate how favorable a given state is to the player of interest. Similarly, action-value network evaluates any decision in the given state (producing an output known as the Q value), allowing us to evaluate the quality of any decision in any state. We thus evaluate the quality of any human decision in any state, as well as the quality of a decision generated by AI, and calculate the gap (i.e., difference) between the two values of decision quality.

Definition of the Human-AI Gap We use notations from the previous literature (Igami, 2020; Silver et al., 2017) to explain our measure more formally. It is defined mostly for the environment of the board game Go, but it can be easily modified in other well-defined dynamic decision-making problems. State space, $|S|$, represents a set of possible states of a game. In a match between human players, a human player faces many states in $|S|$. Given $S \in |S|$, the human player decides on the next move. Thus, a human player making the k^{th} decision observes S_k (state) and decides a_k (action), i.e., a position to place the stone. We simplify the decision rule of human players as follows:

$$a_k^{Human} = \sigma^{Human}(S_k; V(S_k; \theta^{Human}))$$

Human players use their own evaluation parameter (θ^{Human}) to diagnose how advantageous the current state is, $V(S_k; \theta^{Human})$. Based on the evaluation, they apply their own strategy or decision rule (σ^{Human}), ending up with an action, a_k^{Human}. In this decision rule, we abstract away from complex interactions between human players. Instead, we treat a move decision as a single-agent problem in which each human player has to find an optimal decision in the given state to maximize the total reward. Any strategic responses from the opponent human player are subsumed under the transition of the state in our decision rule.

AI programs also map a given state to an action based on their policy network[8]. Although the actual process of AI decision-making is as complex as human decision-making, it can be simplified as follows:

$$a_k^{AI} = \sigma^{AI}(S_k; V(S_k; \theta^{AI})).$$

We need to note that a human and AI facing the same state, S_k, may arrive at different actions (i.e., $a_k^{Human} \neq a_k^{AI}$). This is because the way a human evaluates the state, θ^{Human}, is different from the way AI does, θ^{AI}. That is, a human may be too optimistic or too pessimistic from the perspective of AI. In addition, AI may use a strategy, σ^{AI}, that is distinct from any traditional human strategy, σ^{Human}. AI trained by playing against itself is free from any conventional human strategies, so it will produce actions that will be novel to human players.

Finally, we define the Human-AI Gap as follows:

$$\Delta_k \equiv \overbrace{V(S_{k+1}(a_k^{AI}); \theta^{AI})}^{\text{Quality of a counterfactual AI decision}} - \underbrace{V(S_{k+1}(a_k^{Human}); \theta^{AI})}_{\text{Quality of an actual human decision}}$$

First, a human player takes an action (i.e., makes a decision), a_k^{Human}, after observing a state, S_k. We have AI simulate a counterfactual action (i.e., decision), a_k^{AI}, given the same state S_k; this is the action that the AI itself would have taken if it were in the human player's position. Second, we have the AI quantify the quality of the two actions (i.e., decisions), one from the human player (a_k^{Human}) and the other from the AI (a_k^{AI}), by evaluating the subsequent state. The quality of the counterfactual AI action, $V(S_{k+1}(a_k^{AI}); \theta^{AI})$, represents the "maximum" decision quality that the human player can attain (by choosing the same action as the superhuman AI). Finally, we subtract the human player's decision quality from this "maximum" decision quality to obtain the Human-AI Gap for the move. Put differently, the Human-AI Gap quantifies the difference between the advantage to the human player induced by a counterfactual AI decision and the advantage induced by the player's own decision[9]. If the AI generates a

[8] In most cases, Deep Q network is designed to choose an action with the highest expected action-value. In the game of Go, that process is complemented with Monte Carlo Tree Search and AI programs choose a move with the highest playout number. In addition, the dimension of the state space in the board game Go is very large so AI programs take a few key features of each state as input.

[9] In the game of Go, $V(S_{k+1}(a_k^{Human}); \theta^{AI})$ is similar to Q value (action value network) because the reward is only realized at the end of the game.

higher-quality decision than the human player's, the Human-AI Gap would be positive. If the human player makes the same decision as the AI, then the Human-AI Gap would be zero. Because we use a much superior AI to generate counterfactual decisions, the Human-AI Gap is usually positive. Given this definition of the Human-AI Gap, a systematic decrease in the Human-AI Gap (as we shall see later) indicates that human players make decisions that are closer in quality to decisions of AI.

Data

Our data set spans from January 2014 to March 2020 and consists of two different data sets. The first is data on 30,995 matches between 357 Korean professional players, where we observe match date, the identity of players, and match outcomes. The second is data on 1.3 million move decisions from matches between Korean professional Go players. We scraped publicly available data from the Korean Go Association and other websites[10]. For every move decision by a human player, we simulate the optimal move decision of AI[11] under the same state of the game and compute the Human-AI Gap as explained in the previous section. Thus, we have 1.3 million move decisions by human players, 1.3 million move decisions by AI, and AI's evaluation of each of these decisions.

Table 1 shows the summary statistics of our data. Players in the data were heterogeneous in terms of performance, with the win rates averaging 43%. For each match, two professional players together made an average of 216 move decisions (or 108 move decisions per player). The mean Human-AI Gap of 3.52 percentage points indicates that human players on average made move decisions that were associated with 3.52 percentage points lower win rates as compared with the counterfactual move decisions by the AI program.

Table 1: Summary Statistics

Data 1: Player-level match performance (N = 357)				
	25% quartile	Median	Mean	75% quartile
Winning rate (%)	31	46	43	81
Data 2: Move decisions (N = 1,357,523)				
	25% quartile	Median	Mean	75% quartile
Move counts within a match ($\simeq 2k$)	176	211	216	254
The Human-AI Gap (percentage points)	0.39	1.68	3.52	4.51

Results

Model-free descriptive pattern We calculate the Human-AI Gap (Δ) for every decision made by every human player in our data set. A value of zero means that a human player replicated the AI's decision ($\Delta_k = 0 \iff a_k^{Human} = a_k^{AI}$) and made the optimal move, while a positive value indicates the extent to which AI's decision was superior to that of the human player (i.e., the extent to which the human player's decision quality trailed that of the AI). We present the pattern of Δ_k in Figure 3. The mean Human-AI Gap of each set of 10 moves (e.g., $^{st} - 10^{th}, 11^{th} - 20^{th}$, ...) is plotted across the course of a match. The solid red curve traces the mean Human-AI Gap during the period before AlphaGo's victory over a human Go champion, Lee Sedol (from January 2014 to March 2016); the dotted green curve traces the mean Human-AI Gap during the period between AlphaGo's victory and the release of open-source AI programs, such as Leela Zero (from March 2016 to October 2017); lastly, the dashed blue curve traces the mean Human-AI Gap during the period after the release of the open-source AI programs (from October 2017 to March 2020).

Insights from the pattern of the Human-AI Gap We can trace the inverted U pattern of the Human-AI Gap over the course of a match to gain insights about the human learning from AI. One possible insight is that room for the human learning from AI may have been greatest in the *early middle* stage of the match (Moves 31-60 for each player), as it is where the Human-AI Gap is the greatest. Human players may find that studying AI decisions in this stage improves their game more than studying AI decisions in the early or late stage of the match. Another possible insight is that the human learning from AI in the *middle to late* stage of the match (Moves 51-140 for each player) may have been ineffective. Although human experts have managed to improve their decisions in the *early to early middle* stage of the match (the "Before AI" curve shifted down to "After Open-Source AI" curve for Moves 1-50 in Figure 3), perhaps they could not improve their decisions in later stages of the match despite much effort (the "After Open-Source AI" curve overlaps the "Before AI" curve for Moves 51-140 in Figure 3). Human experts may have concluded that improving decisions in later stages of the match (Moves 51-140 for each player) was more difficult than improving decisions in the *early to early middle* stage (Moves 1-50). Improving decisions in the *middle* stage (Moves 51-90 for each player) may have been especially hard, perhaps due to intractable complexity and lack of similarity from one match to another, both of which may have prevented discovery or learning of novel strategies. If so, human experts may instead have doubled down and focused their effort even more on improving decisions in the *early to early middle* stage (Moves 1-50).

Human learning in the early stage of the game More important than the inverted U pattern are downward shifts in the Human-AI Gap (for Moves 1-50). Interestingly, the Human-AI Gap decreased only a little bit after human players could observe AlphaGo's actions, as evidenced by

[10] Not every match has been saved with a detailed record of move decisions, but we collected the historical data from multiple sources to get more complete history. The data containing match results spans from 2012 to 2020. The data containing move decisions spans from 2014 to 2020.

[11] We use an AI program called *Leela Zero* to analyze our data. We use GPU provided by Google Colab Pro (P100) in our simulation.

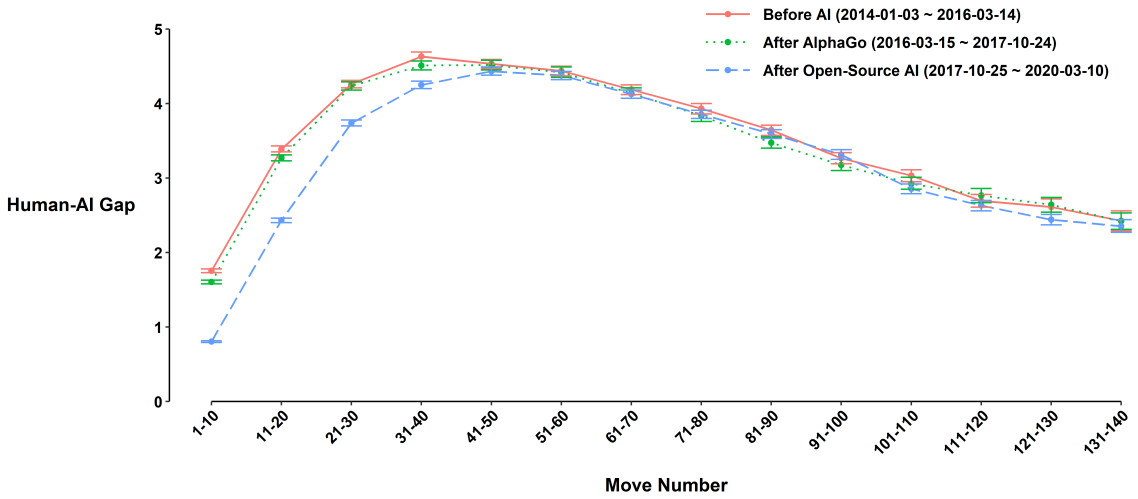


Figure 3: **Model-free Patterns of the Human-AI Gap over the Course of Match (Mean of Each 10 Moves).** the Human-AI Gap decreases ever so slightly after AlphaGo's *actions* are released: the dotted green curve is positioned slightly below the solid red curve. After open-source AI programs are released and *reasoning processes* of AI become observable, however, the Human-AI Gap decreases by a much greater margin, indicating a marked increase in human decision quality. This decrease in the Human-AI Gap occurs mostly for the *early to early middle* stage of the match (Moves 1-50). The error bars indicate 95% confidence intervals around the means for each 10 moves.

a barely noticeable downward shift from the "Before AI" curve to "After AlphaGo" curve in Figure 3. In contrast, the Human-AI Gap dropped markedly after open-source AI programs were released, as evidenced by a larger downward shift from "After AlphaGo" curve to "After Open-Source AI" curve in Figure 3.

Constructing a player-month level data set So far, we have not taken into account differences among players. Now, we construct a data set at the player-month level (Δ_{it}) as follows:

$$\Delta_{it} = \frac{1}{n_{it}} \frac{1}{K} \sum_{j=1}^{n_{it}} \sum_{k=1}^{K} \Delta_{jk}$$

where Δ_{it} denotes player i's mean Human-AI Gap in month t; n_{it} denotes the number of matches player i plays in month t; K denotes the total number of move decisions examined for each player in each match; k denotes player i's kth move decision within match j; and Δ_{jk} denotes the Human-AI Gap for player i's kth move decision within match j. Using a panel structure, we investigate whether human decision quality indeed increased more after the release of open-source AI programs as compared with after AlphaGo's victory. Because the decrease in the Human-AI Gap was concentrated in the *early to early middle* stage of the game (Moves 1-50), and because no such decrease was readily observable for Moves 51-140, we focus our attention on the first 50 moves by each player within each match (i.e., $K = 50$) to investigate human decision quality over time in the following section.

Players with versus without access to AI We investigate how access to reasoning processes of AI (i.e., access to the open-source AI programs like Leela Zero) affects the human learning from AI. We do so by leveraging mandatory military service in South Korea[12] as a natural experiment. First, we obtained an official record of Korean players' military service history and used it to split the players into two groups: those who were serving in the military and therefore did not have access to AI programs for at least 6 months; and those who were not serving in the military and had full access to AI programs. Next, we constructed a player-month level data set (as explained in the previous section), separately for each of the two groups of players. Then, we ran the following regression, separately for each of the two groups of players:

$$\Delta_{it} = \alpha_i + \tau_t + \varepsilon_{it}$$

where Δ_{it} denotes player i's mean Human-AI gap in month t; α_i denotes a player fixed effect; τ_t is a fixed effect of individual months (from January 2014 to March 2020), and ε_{it} is an error term. In Figure 4, we plot the estimated fixed effects of individual months, τ_t, for each of the two groups of players. The two vertical lines mark the introduction of the two AI programs: the vertical line on the left marks March

[12] South Korean male citizens are required to serve in the military for 18-24 months before age 29, which forces the military-serving players not to be able to participate in Go matches more than once a month and to be away from recent trends in AI programs related to Go. Specifically, most of them are expected to be confined in their military base but they get short-term leave every other month. We confirm from the data that players serving in the military can participate in a tournament once a month at most. They do not have enough time nor a high-performance computer to self-teach unfamiliar tactics, strategies, or insights discovered by AI programs.

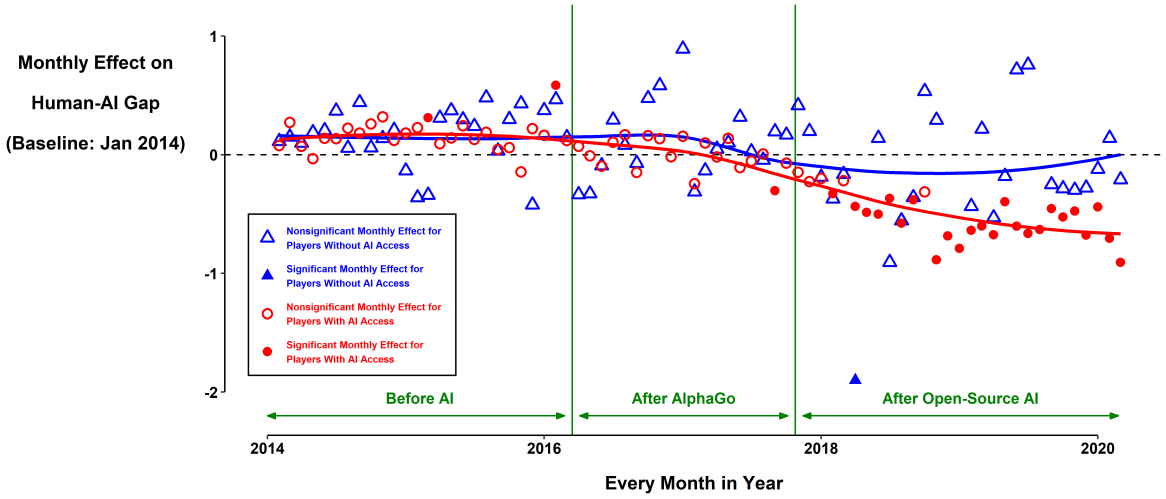


Figure 4: **Statistical Test on the Human-AI Gap** This plot shows estimated monthly fixed effects on the Human-AI Gap (Δ_{it}) for players with access to the AI programs (denoted by red circles) and for players without the access (denoted by blue triangles). The following regression equation was used to estimate the time trends: $\Delta_{it} = \alpha_i + \tau_t + \varepsilon_{it}$, where α_i is a player fixed effect, τ_t is a time (i.e., month) fixed effect, and ε_{it} is an error term. This plot demonstrates that AlphaGo, despite its superior performance against human, does not help human players to make better decisions. Human players start to make better decisions after the release of the open-source AI. This finding highlights the importance of access to *reasoning process* of AI.

15, 2016, the date on which AlphaGo defeated the human Go champion, Lee Sedol; and the vertical line on the right marks October 25, 2017, the date on which the open-source AI program Leela Zero was publicly released (which in turn was followed by releases of similar AI programs and education tools). For each group of players, solid shapes denote significant monthly fixed effects, while hollow shapes denote nonsignificant monthly fixed effects. Players who had access to the open-source AI programs (e.g., Leela Zero) significantly reduced the Human-AI Gap in the months following the introduction of Leela Zero: the monthly fixed effects were mostly nonsignificant (hollow red circles) before the release of Leela Zero, but they become mostly significant afterwards (solid red circles). In contrast, players in the military who had difficulty in getting access to the AI programs did not show such reduction in the Human-AI Gap: the monthly fixed effects for this group are nonsignificant throughout the whole period (hollow blue triangles), except for one month, April 2018 (one solid blue triangle). In summary, for players with access to the AI programs, the Human-AI Gap starts to decrease significantly after the release of open-source AI programs (the red smooth line tracing monthly fixed effects slopes downward noticeably after the second vertical line). In contrast, for players without access to the AI programs, the Human-AI Gap remains relatively flat (the blue smooth line does not noticeably slopes downward). This pattern of monthly effects suggests that AlphaGo, despite its superior performance against humans, did not help human players make better decisions. Human players learned from AI after the open-source AI programs revealed the reasoning processes of AI, rather than after AlphaGo revealed actions of AI[13]. This finding highlights the importance of gaining access to *reasoning process* of AI to promote the human learning from AI.

Conclusion

As AI technology advances, we are likely to witness AI outperforming humans in many decision-making domains other than Go. Go is an interesting decision-making domain to study the human learning from AI because it has a unique combination of two notable features: (1) AI produces decisions through a complex process that humans cannot fully comprehend; and (2) the link between each AI move decision and the ultimate outcome of the match (win vs. loss) may be quite unclear to humans, though AI can make better sense of the unclear link. Despite such complexity and ambiguity inherent in decision-making by AI in Go, our results suggest that humans can learn from AI and make better decisions like AI, provided that they have access to the *reasoning processes* of AI. Professional Go players did not make better decisions after observing mere *actions* of the AI that defeated a human Go champion. Rather, the human experts started making better decisions only after they gained insight into *reasoning processes* of the open-source AI programs. We found further evidence of this when we compared human players who had versus did not have access to the AI programs. Even as AI surpasses humans in decision-making, humans may find it difficult to learn from AI and make better decisions themselves, unless they can have an inside look at how AI makes its decisions.

[13]We verify the finding by doing a Difference-in-Difference estimation in our online appendix (https://osf.io/skuy5)

Acknowledgments

We appreciate feedback from Kosuke Uetake, Jiwoong Shin, K Sudhir, Elisa Celis, John Rust, Vineet Kumar, Ian Weaver and participants at the 2020 Marketing Science conference and Yale Quantitative Marketing seminar for their valuable and constructive comments. We gratefully acknowledge the support from Korean Go Association to get the data.

References

Biswas, T. T. (2015). Measuring intrinsic quality of human decisions. In *International conference on algorithmic decisiontheory* (pp. 567–572).

Igami, M. (2020). Artificial intelligence as structural estimation: Deep blue, bonanza, and alphago. *The Econometrics Journal*.

Regan, K. W., Biswas, T., & Zhou, J. (2014). Human and computer preferences at chess. In *Workshops at the twenty-eighth aaai conference on artificial intelligence.*

Resnick, L. (2018). *Knowing, learning, and instruction: Essays in honor of robert glaser*. Routledge.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... others (2017). Mastering the game of go without human knowledge. *Nature*, *550*(7676), 354–359.

Stafford, T., & Dewar, M. (2014). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychological science*, *25*(2), 511–518.

Strittmatter, A., Sunde, U., & Zegners, D. (2020). Life cycle patterns of cognitive performance over the long run. *Proceedings of the National Academy of Sciences*, *117*(44), 27255–27261.

Tsividis, P. A., Pouncy, T., Xu, J. L., Tenenbaum, J. B., & Gershman, S. J. (2017). Human learning in atari. *2017 AAAI Spring Symposium*.