UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Recursive Adversarial Reasoning in the Rock, Paper, Scissors Game

Permalink

https://escholarship.org/uc/item/4sp2722s

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 42(0)

Authors

Brockbank, Erik

Vul, Edward

Publication Date

2020

Peer reviewed

Recursive Adversarial Reasoning in the Rock, Paper, Scissors Game

Erik Brockbank (ebrockbank@ucsd.edu)
Department of Psychology, 9500 Gilman Drive
La Jolla, CA 92093-109 USA

Edward Vul (evul@ucsd.edu)
Department of Psychology, 9500 Gilman Drive # 109
La Jolla, CA 92093-109 USA

Abstract

In this study, we investigate people's ability to predict and adapt to the behavior of others in order to make plans of their own, a cornerstone of cooperative and competitive behavior. Participants played 300 rounds of rock, paper, scissors against another human player. We investigate the degree to which participants are able to identify patterns in their opponent's behavior in order to exploit them in subsequent rounds. We find strong evidence that participants exploit their opponents over the course of 300 rounds, suggesting that people identify dependencies in their opponent's move choices during the game. Nonetheless, analysis of dependencies across participant move choices reveals that people exhibit a number of regularities in their own moves. Based on these dependencies, we argue that participants are far from optimal in their exploiting, suggesting that there are substantial constraints on people's ability to identify and adapt to patterned opponent behavior across repeated interactions.

Keywords: adaptive reasoning; adversarial reasoning; non-cooperative games; rock paper scissors

Introduction

Human conflict and coordination relies on our ability to reason about and predict the behavior of others in order to make plans of our own. What kind of cognitive processes underlie this unique ability? In the current experiment, we investigate this question in an adversarial setting which highlights the ways in which people detect and exploit patterns in an opponent's behavior across repeated interactions. We use the simple game of rock, paper, scissors (RPS) to evaluate the level at which people are able to flexibly reason about their opponents in order to maximize their own gain.

RPS is an example of a zero-sum, *cyclic dominance* game (Morgenstern & Neumann, 1953); each move choice can be beaten by another move choice, meaning no move is preferable to any other without some expectation about what the opponent will do. This forces players interacting over repeated rounds to continually shuffle between moves in a way that avoids exploitation by their opponent, while seeking any pattern in their opponent's choices that will give them an upper hand. Because the game can be played over many rounds successively and with the same opponent, it allows participants to detect and adapt to potentially complex patterns in their opponent's behavior. Additionally, the game has simple rules and a constrained space of choices, such that better performance by one individual in a dyad will not likely be a result of particular "expertise" but rather a result of better reasoning about dependencies in their opponent's move choices.

For these reasons, RPS provides an ideal setting for asking questions about people's ability to reason about and predict the behavior of others over repeated interactions.

Despite these benefits, little work has examined the cognitive underpinnings of the game and the inferences involved in playing it successfully. Wang, Xu, and Zhou (2014) explore whether participants playing RPS in shuffled groups show behavior closer to the Nash Equilibrium strategy of random play (Nash, 1950) or the cyclic behavior predicted by models of evolutionary stable strategies (Smith, 1982). Importantly for the present work, they suggest that individual decisions in the game are best characterized according to a "win-stay, lose-shift" conditional strategy. However, because players' opponents were continually shuffled throughout the game, these results may not predict the ways people reason about and adapt to the behavior of a stable opponent.

Subsequent studies have found further evidence that people's decisions in repeated RPS games are consistent with win-stay, lose-shift behavior to varying degrees (e.g., Dyson, Wilbiks, Sandhu, Papanicolaou, and Lintag (2016); see Dyson (2019) for review), but these have typically involved humans playing against randomized bot opponents, which also leaves open the question of whether such strategies are effective against an adaptive human opponent, or simply represent a sort of fallback against opponents that are unexploitable. A number of studies have explored people's ability to identify and exploit *strategic* bot opponents (Stöttinger, Filipowicz, Danckert, & Anderson, 2014), finding that people can adapt to a range of different regularities in opponent strategy. In the present work, we explore what happens when people are paired against other humans who may be equally attentive to particular strategies or behavioral regularities. While a handful of studies have paired participants against stable human opponents in the rock, paper, scissors game (e.g., Baek et al. (2013)), these studies have used the game to address questions unrelated to the present work, such as whether people imitate gesture subconsciously (Cook, Bird, Lünser, Huck, & Heyes, 2012; Aczel, Bago, & Foldes, 2012).

Within the broader class of *mixed strategy equilibrium* games, of which RPS is a single example, research involving stable human dyads has mostly examined whether people are able to adopt the Nash Equilibrium strategy of random play over many rounds (for a thorough review, see Camerer (2011)) but do not specifically explore people's ability to

adapt to dependencies exhibited by their opponents. More broadly, cyclic games similar to RPS that can be scaled to larger groups of players have been used to examine large scale collective behaviors (Frey & Goldstone, 2013); however, these findings are mostly concerned with the cognitive underpinnings of group-level behavior rather than adaptive reasoning about an individual opponent. Meanwhile, research focused on recursive reasoning about other individuals over repeated interactions has addressed a number of challenges, from pragmatics and convention formation (Hawkins, Goodman, & Goldstone, 2019), to lying and deception (Oey, Schachner, & Vul, 2019), but does not make specific predictions about how participants will behave in an adversarial setting like the RPS game.

Experiment

In this experiment, we use the simple game of rock, paper, scissors (RPS) to investigate the degree to which people can detect patterns in their opponent's behavior and exploit these patterns in an adversarial setting.

Participants

Participants were 128 college students who received course credit for their participation. Participants were assigned to stable dyads during the experiment (except in two cases where odd numbers of participants precluded dyad formation); of 62 dyads, four were removed due to technical issues which prevented their completion of the task, leaving 58 dyads with complete data.

Procedure

Participants began by clicking through a set of instructions introducing the game of rock, paper, scissors and noting that they would be playing against a human opponent. Upon completion of the instructions, participants entered a "waiting room" where they were matched up into dyads on a first in, first out basis. Once a participant had been paired up with another participant, they began the RPS game, which consisted of 300 rounds against the opponent they had been matched with. In each round, they were shown a set of clickable "cards" with rock, paper, and scissors icons and instructed to choose a move. They were given 10 seconds to choose their move each round[1]. Once a participant had chosen a move, they could not change their selection. The first player to choose a move was taken to a screen that highlighted their move and instructed them to wait while their opponent selected a card. After both participants had chosen a move, they were taken to a screen indicating each player's move, the results of the round, and the points each player received for that round. Participants were given 3 points for a win, 0 points for a tie, and –1 points for a loss. Participants could view the results of the round for as long as they wanted before clicking a button to proceed to the next round; only after both participants had clicked "Continue" did the next round begin (the participant who clicked first remained on the results screen with a message indicating that the next round would begin as
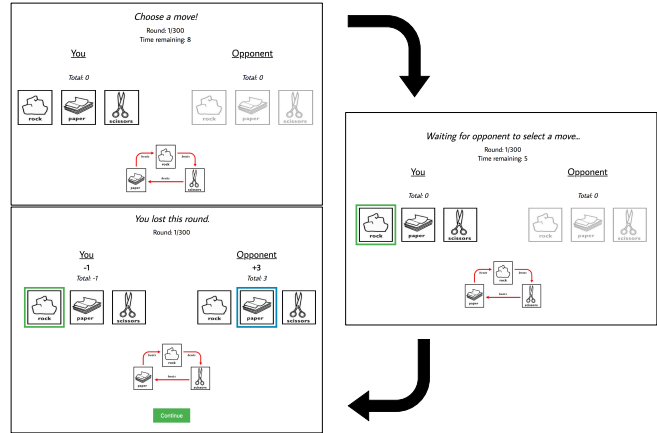

Figure 1: The three stages of each rock, paper, scissors round. At top left, selecting a move. At middle right, waiting for opponent after selecting a move (this screen is bypassed if a player is the last to choose a move). At bottom left, the results of a round after both players have selected a move.

soon as the other player was ready). A graphic illustrating the rules (which card beat which) and a tally of rounds completed and total points for each player were visible throughout the game. Figure 1 illustrates the phases of each round.

Upon completion of all 300 rounds, participants were directed to a post-game questionnaire. Most participants completed the experiment in under 30 minutes (mean time from round one to round 300: 966s, sd: 347s). The experiment was coded in node.js using socket.io, a popular library for synchronous game play in web browsers[2](Hawkins, 2015).

Results

What results might be *expected* for dyads playing 300 rounds of rock, paper, scissors against a stable opponent? The Nash Equilibrium strategy (Nash, 1950) in repeated play is to choose moves randomly; any non-random play necessarily produces patterns which could be exploited by an optimal opponent. Prior research has shown that characteristic patterns of *subjective randomness*, which may be highly predictable, emerge even in adversarial settings where such predictability is a disadvantage (Budescu & Rapoport, 1994). This sets up a natural challenge for a player engaged with the same opponent over many rounds to detect and respond to any patterns in their opponent's play. We investigate whether participants show evidence of such reasoning about their opponents.

How much are people exploiting?

First, we look at game results across dyads to determine whether participants in our sample were able to capitalize on

---

[1]Participants typically chose a move after no more than a few seconds: average response time across participants was 1,228ms (sd: 642ms).
[2]The code for this experiment, as well as the analyses presented here, are available at: https://github.com/erik-brockbank/rps

dependencies that their opponents exhibited. If participants were playing randomly and not attempting to exploit their opponents, they should have an equal number of wins, ties, and losses on average. However, to the degree that a single player in a dyad is able to anticipate their opponent's move choices, this should allow them to obtain a higher win count over the course of 300 rounds. Figure 2 shows the empirical distribution of win count *differentials* across all 58 dyads: we take the absolute value of one player's win count subtracted from the other's in each dyad. Overlaid on the same figure is the distribution of win count differentials that would be expected if dyads were playing randomly. This null distribution is generated based on 10,000 samples and approximates the absolute value of a Gaussian centered at zero.

It's clear in Figure 2 that the empirical distribution of win count differentials has a heavier tail than would be expected under random play. This density at larger win count differentials suggests that many participants were able to exploit their opponents, thus obtaining more wins over 300 rounds than would be expected by chance. The empirical distribution has a mean of 19.26 (SEM: 1.71), compared to the null distribution's mean of 11.3; the difference between empirical win counts and the proportions under the sampled null distribution is highly significant (χ^2 (5) = 133.27, $p < 0.001$)[3]. Additionally, when we truncate the empirical data to remove the top 10%, 20%, and 30% of win count differentials, the chi-squared comparison with the sampled null data remains significant. Therefore the difference between the null and empirical win count differentials is not simply driven by a few outlier dyads.

The fact that people seem to be exploiting patterns in their opponents' behavior as shown in Figure 2 raises a question about the level at which participants are responsive to dependencies in their opponent's play. One possibility is that participants detect patterns in their opponents' moves from one round to the next, i.e., transient dependencies that they are able to leverage for short periods. A second possibility is that players are not responsive to their opponents at such a fine-grained level, but rather are sensitive to aggregate patterns in their opponents' move choices which allow them to accrue a modest but stable advantage.

If players are responsive to transient patterns in their opponents' moves and are able to exploit these patterns in short stretches, then we would expect these players to win in *streaks* throughout the 300 rounds. These winning streaks would constitute evidence of one player making a series of correct predictions about their opponent's moves before their opponent is able to notice and perhaps predict what the original exploiter is doing. On the other hand, if players are obtaining higher-than-expected win count differentials simply by tracking more stationary dependencies in their opponents' moves, then we should not expect winning or losing streaks so much as a pattern of uncorrelated outcomes which lead one

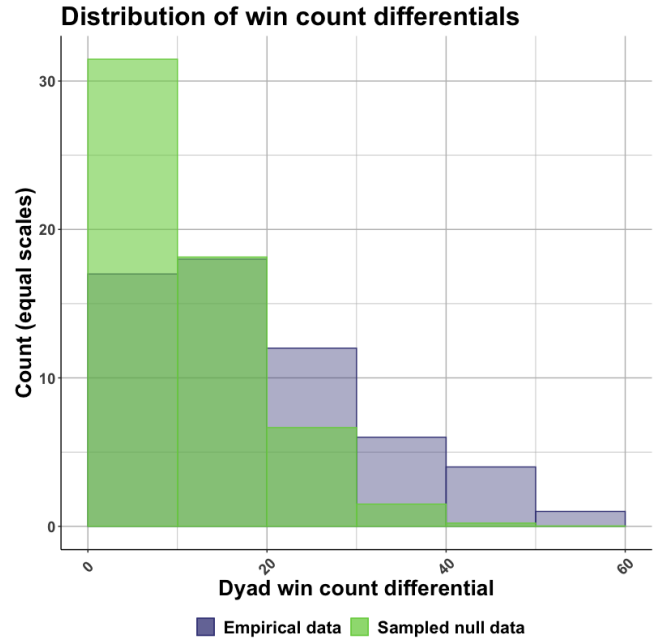**Distribution of win count differentials**

Figure 2: Empirical and null distributions of win count differentials within dyads. The null distribution is based on 10,000 samples but scaled to match the empirical counts. The empirical distribution has a heavier tail, suggesting that many players successfully exploited their opponents.

player to have more wins by the end of the game.

To investigate whether participants are able to exploit transient patterns in their opponents' move choices, we look at the auto-correlation of round outcomes over increasing stretches of game rounds. If a player is more likely to win after a round in which she has just won, then we expect an auto-correlation of outcomes at a *lag* of one round. Figure 3 shows the auto-correlation of round outcomes over increasing game round lags. Each of the green points represents a single dyad and the dark blue is the average across dyads. The dashed lines indicate a 95% confidence interval around a mean auto-correlation of zero. Though there is one dyad which had a high degree of outcome auto-correlation and some that approached meaningful correlations at a lag of one or two rounds, we find very little dependency on previous outcomes in the full set of dyads. The data are similar when we examine just those participants who had the highest win count differentials (i.e., there doesn't seem to be a relationship between beating one's opponent more over 300 rounds and having a substantial auto-correlation of outcomes). This suggests that people's ability to exploit their opponents may not come from detecting short but predictable patterns in their opponents' moves, but rather from exploiting overall statistics in their opponents' choice probabilities across many rounds.

A natural question that arises from these results is whether participants were able to detect short-term patterns in their opponents' moves that simply fail to meet the threshold for detection with our auto-correlation measure. To assess the

---

[3]Chi-squared comparisons remain significant when win count differentials are binned by all values between 5 and 20.
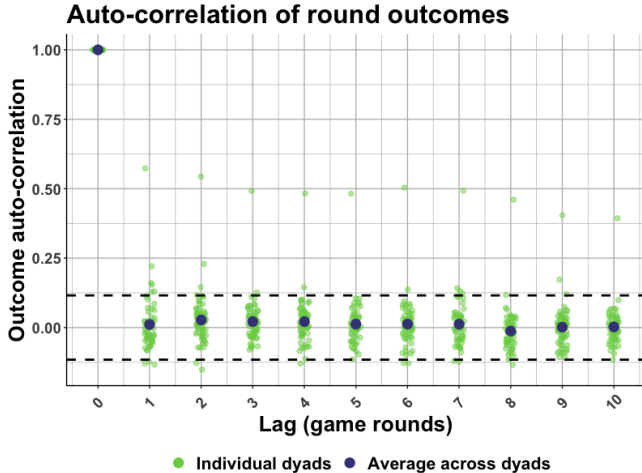
Auto-correlation of round outcomes

Figure 3: Auto-correlation of round outcomes for all dyads. Green points are individual dyads and blue points are the average across dyads. The dashed lines indicate 95% confidence intervals around a mean of zero. Though individual dyads show some evidence of correlated outcomes at a lag of one or two rounds, there is little evidence in aggregate of participants exploiting their opponents based on transient move patterns enough to generate a significant auto-correlation.

sensitivity of the auto-correlation, we simulated 300-round games of RPS with 1,000 samples, where the outcomes were uniformly selected but then a subset of those outcomes were replaced with a k-length streak of wins and a k-length streak of losses. Though empirical dyad results are unlikely to have streaks distributed in this way, it provides an idealized baseline for detecting auto-correlations in win-loss patterns. We find that when streaks make up 10% of game rounds above chance (i.e., a 15-round win streak and a 15-round loss streak), lag-one and lag-two auto-correlation is detected with greater than 50% power for individual dyads. With streaks of 20 wins and 20 losses, auto-correlations at a lag of one and two are detected above 80% power. Finally, with streaks composing 20% of game rounds above chance, power for individual dyads is greater than 99% for lags of one, two, and three and greater than 90% for lags all the way out to 10. This suggests that participants who are exploiting their opponents would need as few as 30 streaky wins and losses above what is expected by chance over the course of 300 rounds to produce significant auto-correlations at a lag of one; the longer these streaks, the more likely they achieve significant auto-correlations at greater lags. Broadly, the auto-correlation measure is reasonably sensitive to streaky outcomes across 300 rounds of rock, paper, scissors. Therefore, it seems unlikely that the distribution of win count differentials seen in our empirical data is fully accounted for by transient, streaky exploitation of a player's opponent.

How exploitable are people?

Our initial results show that participants exploit dependencies in their opponent's move choices more than would be expected if people were playing randomly, leading to higher rates of large win count differentials. However, the low auto-correlation of outcomes across each dyad's 300 rounds suggests that people may simply be exploiting stable tendencies in their opponents' move choices rather than adapting to flexible policies over the course of the game. Does this behavior reflect participants extracting as much of a pattern as their opponents reliably display? Or do people have rich dependencies in their move choices that their opponents fail to detect?

To address whether participants are exploiting *optimally*, we examine how exploitable they are to begin with. Based on the statistical dependencies that a player exhibits in their move choices, an optimal opponent can choose moves to maximize the *expected win count differential* obtained by leveraging those dependencies. Intuitively, if a player tends to favor "rock" over "paper" and "scissors", their opponent has a higher expected win count differential over many rounds by choosing "paper" more frequently. Critically, the maximum expected win count differential for a given player's opponent depends on which statistical dependency the opponent aims to exploit: the player's move probabilities given their previous move may yield a different expected win count differential than move probabilities given their previous five moves. Whichever of these is more predictive will produce a higher maximum expected win count differential each round.

Formally, let a dependency structure $\mathbf{X}$ represent the set of possible prior events that influence a player's move probabilities; for instance, consider a previous-move dependency in which a player's (very exploitable) opponent chooses rock, paper, or scissors based simply on whether their last move was rock, paper, or scissors: $\mathbf{X}$ designates the full set of possible previous move states. Let $p(i \mid x)$ be the probability that the player's opponent chooses a particular move $i \in \{\text{"rock"}, \text{"paper"}, \text{"scissors"}\}$ following some history of play $x \in \mathbf{X}$. If $\mathbf{X} = \{\}$ then $p(i \mid x)$ is simply the average probability of the opponent choosing each move; if $\mathbf{X}$ encodes more complicated dependency structures, e.g., the opponent's previous move, then $p(i \mid x)$ could indicate the probability of a move i given the opponent's previous move (in this case x would indicate whether the previous move was rock, paper, or scissors). Next, let $v(i, j) \in \{-1, 0, 1\}$ be the outcome of playing a particular move j against the opponent's move i: increasing the player's win count by 1, decreasing by 1, or tying for a change of 0. A player has an expected win count differential for a particular move j which is defined over all possible moves i that their opponent might play. We can define a player's expected win count differential by choosing move j in a game round following a particular history x by averaging over all possible opponent moves i: $\sum_i p(i \mid x) \cdot v(i, j)$. For each possible prior event history $x \in \mathbf{X}$ (e.g., each possible opponent previous move), a player has a move j_x^* that maximizes the expected win count differential over all possible

opponent moves i:

$$j_x^* = \arg\max_j \left(\sum_i p(i \mid x) \cdot v(i, j) \right)$$

. Using this expression, we compute the expected win count differential for optimal play each round by averaging over all possible history outcomes $x \in \mathbf{X}$ that might occur in a particular round:

$$\frac{1}{|\mathbf{X}|} \sum_{x \in \mathbf{X}} \sum_i p(i \mid x) \cdot v(i, j_x^*)$$

. This is the maximum expected win count differential a player obtains each round by assuming the opponent will choose their moves according to the dependency encoded in $\mathbf{X}$. By multiplying this by 300, we get a player's expected win count differential with optimal play (against dependency $\mathbf{X}$) for a full game of 300 rounds. Importantly, this provides a measure of how much participants could plausibly be exploited by a given dependency in their move choices: the more a player exhibited a particular dependency in their moves, the higher the maximum expected win count differential against that dependency.

To understand the degree to which participants exhibited patterned regularities in their behavior, we compute maximum expected win count differentials for each participant based on various dependencies in their move choices. Figure 4 shows empirical win count differentials, the sampled null win count differential described previously, and maximum expected win count differentials (averaged across participants) for eight unique statistical dependencies in participant behavior. We review each of these below.

Expected win count differentials for move distributions The most intuitive dependency people might exhibit is in their distribution over *moves* given various combinations of their own and their opponent's previous moves. In Figure 4, we examine maximum expected win count differential based on five dependencies players exhibit in their move distributions (the other three dependencies in Figure 4, which are not reflected in move distributions, are discussed in the next section). If players display a bias towards particular moves given any of these dependencies, this will increase the maximum expected win count differential based on that dependency.

`Choice baserate (R/P/S)` indicates each player's overall distribution of moves.

`Choice given player's prior choice` indicates each player's distribution of moves given their own previous move.

`Choice given opponent's prior choice` indicates each player's distribution of moves given their *opponent's* previous move.

`Choice given player's prior two choices` indicates each player's distribution of moves given their own previous *two* moves.

`Choice given player's prior choice & opponent's prior choice` indicates each player's
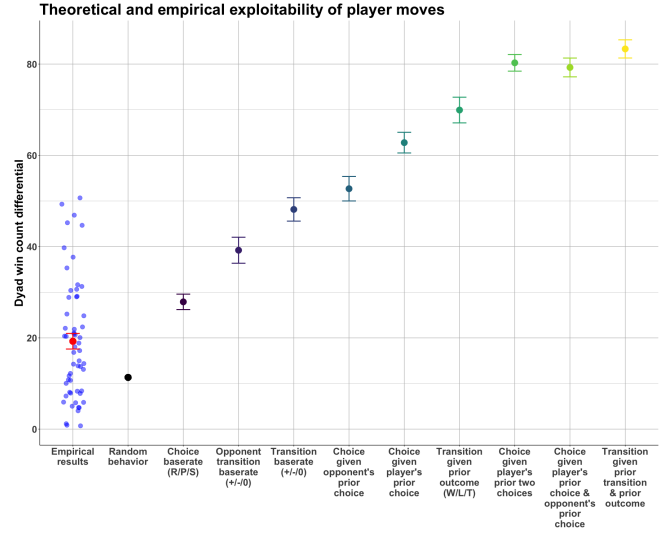

Figure 4: Win count differentials for empirical data, sampled null data representing random behavior, and *maximum expected win count differentials* exploiting various dependencies in participant move choices and transitions. Error bars indicate one SEM.

distribution of moves given the combination of their own previous move and their opponent's previous move.

Together, these dependencies represent a range of ways in which a player might be more predictable in virtue of their move choices being conditionally dependent on combinations of prior moves. Several notable findings emerge from the expected win count differentials for each of these dependencies in Figure 4. First, the expected win count differential using each player's overall distribution of moves (`Choice baserate (R/P/S)`) is already larger than the empirical average win count differential. This suggests that noticing a dependency on overall move probabilities may have been sufficient for winning players to exploit their opponents in some of the dyads which had higher than expected win count differentials. Next, it's clear in Figure 4 that while participants do exhibit some dependency on their opponent's previous move (`Choice given opponent's prior choice`), they exhibit a greater dependency on their *own* previous move (`Choice given player's prior choice`). This suggests that while participants are likely attending to what their opponents do at some level, their move choices may be based far more on their own previous moves. Finally, we note that among the best predictors of a player's next move in our data are the player's previous two moves in combination (`Choice given player's prior two choices`) and the player's previous move along with the opponent's previous move (`Choice given player's prior choice & opponent's prior choice`)[4]. The large expected win count differential for these two dependencies suggests that a player who tracks their opponent's previous two moves or their opponent's previous move and their own previous move

Figure 5: The "transitions" possible between a player's previous move and their next move. These can also be viewed as transitions relative to the *opponent's* previous move.

has a significant opportunity to exploit their opponent, though it is unclear to what degree participants in the game used this information to obtain higher win count differentials.

Expected win count differentials for transition distributions While the intuition that people show statistical dependency on previous moves in their own move choice is straightforward, previous work has argued that people playing repeated games of RPS show evidence of more sophisticated strategies, such as *win-stay, lose-shift* (Wang et al., 2014), a policy that can be used to approximate Bayesian inference (Bonawitz, Denison, Gopnik, & Griffiths, 2014). This would involve playing the same move after a win but switching in a reliable *direction* after a loss. Move choices can be thought of as having directionality by considering transitions between moves: a "positive" shift involves playing paper after rock, scissors after paper, and rock after scissors while a "negative" shift moves in the opposite direction: see Figure 5 for a full schematic of RPS transitions. A sophisticated player might follow a pattern of transitions given previous outcomes or transitions, rather than having a stationary dependency on their previous moves. Critically, transitions can be relative to a player's own previous move (e.g., playing the move that will beat what they just played) or their *opponent's* previous move (e.g., playing the move that will beat what their opponent just played). We consider both forms of transition below.

In Figure 4, we examine three possible dependencies which might impact people's transition probabilities from one move to the next[5]. As in the previous section, the more predictable a player's move transitions are based on these dependencies, the higher the maximum expected win count differential an optimal opponent obtains by leveraging these dependencies.

Opponent transition baserate (+/-/0) indicates the overall distribution of transitions, i.e., how often players shift up (+), down (−), or stay (0) relative to the *opponent's* previous move.

Transition baserate (+/-/0) indicates the overall distribution of transitions, i.e., how often players shift up (+), down (−), or stay (0) relative to their *own* previous move.

Transition given prior outcome (W/L/T) indicates the distribution of transitions (+, −, 0) given a player's previous outcome (win, loss, tie). A win-stay, lose-shift dependency would be exhibited at this level, but many different outcome-dependent transitions are possible.

Transition given prior transition & prior outcome indicates the distribution of transitions given a player's previous outcome and previous transition. As an example, a player might be more likely to shift up (+ transition) if they won on the previous round by shifting up.

Expected win count differentials for transition dependencies, compared to move dependencies, reveal a number of interesting results. First, we note that the distribution of overall transitions (Transition baserate (+/-/0)) yields less exploitable signal than the distribution of moves given a player's previous move, despite the fact that both of these rely on the same information. This suggests that people's dependencies on their previous moves do not boil down to simple transition dependencies but are more complex. As with the move dependencies, people are more exploitable on the basis of transitions made relative to their own previous move (Transition baserate (+/-/0)) than relative to their opponent's previous move (Opponent transition baserate (+/-/0)), once again suggesting that people may be more likely to draw on their own past history when choosing moves rather than their opponent's. Next, we note that the expected win count differential for a player's transition given their previous outcome (Transition given prior outcome (W/L/T)) is larger than the dependencies participants exhibit on their previous move or their opponent's previous move. The fact that participant transitions are fairly predictable based on their previous round outcome provides suggestive evidence in line with prior work that a strategy like win-stay, lose-shift may be widely used, whether intentionally or unwittingly (Wang et al., 2014). Finally, Figure 4 shows the largest expected win count differential for the distribution of transitions given the player's previous outcome and previous transition (Transition given prior transition & prior outcome). This may indicate that people's tendency towards policies like win-stay-lose-shift are subservient to more complex (but ultimately more predictable) policies based on previous outcome *and* previous transition. While it is unclear from our empirical dyad results whether participants exploited such dependencies in their opponents (even partially or transiently), our data suggest that people are highly exploitable at this level of dependency.

---

[4]The expected win count differential for the opponent's previous two moves is lower than these two dependencies so we have omitted it from Figure 4.

[5]As in the previous section, we omit dependencies from these results which do not have theoretically interesting results.

Discussion

In this experiment, we investigate people's ability to plan and reason about others in an adversarial setting using the simple game of rock, paper, scissors, a game that makes it possible for players to detect and exploit a variety of patterns in their opponent's behavior. We find strong evidence that participants did indeed exploit their opponents beyond what would be expected by chance or random play. However, we find little evidence that such exploitation occurs at the level of transient patterns in opponent moves. Rather, outcome autocorrelations are more consistent with players tracking high level dependencies in their opponents' moves. This raises the question of whether participant choices reflect optimal exploitation of their opponents. By comparing the empirical distribution of win count differentials across dyads to the *maximum expected* win count differential for various statistical dependencies in people's move choices, we find that optimal exploitation of people's move selections based on their previous moves and outcomes would produce much larger win count differentials than we see empirically.

However, this does not imply that people entirely fail to exploit the dependencies in their opponents' move choices that we examine here. Instead, it suggests that people's ability to detect regularities in their opponents' behavior is limited to a smaller range of patterns or biases in their opponents' move choices. This raises several intriguing questions which we hope to explore in future work. First, what *are* the set of patterns and biases that a person can optimally exploit in their opponent's move choices over many rounds of RPS? Second, what sorts of cognitive limitations prevent people from learning more complex dependencies exhibited by their opponents? Is such learning simply constrained by working memory for past moves and outcomes? Or is it that people lack the ability to abstract over previous moves and outcomes in a way that would allow them to recognize patterns such as win-stay, lose-shift? Or, is it simply that participants are fairly optimal learners but they do not get enough data over the course of 300 rounds of RPS to strongly differentiate between various strategies their opponents might be using?

Future work exploring these questions has ramifications far beyond the simple game of rock, paper, scissors. Our ability to compete and coordinate with others relies on a fundamental ability to *predict* what others will do next in order to plan our own actions accordingly. In many cases, this involves drawing on knowledge from past interactions and building abstractions over the sorts of strategies or motivations that governed past actions to make optimal predictions. A better characterization of the pattern recognition, abstraction, and learning processes underlying this ability will allow for an increased understanding of how humans so effortlessly work alongside each other in many domains.

References

Aczel, B., Bago, B., & Foldes, A. (2012). Is there evidence for automatic imitation in a strategic context? *Proceedings of the Royal Society B: Biological Sciences*, *279*(1741), 3231–3233.

Baek, K., Kim, Y. T., Kim, M., Choi, Y., Lee, M., Lee, K., ... Jeong, J. (2013). Response randomization of one-and two-person rock–paper–scissors games in individuals with schizophrenia. *Psychiatry research*, *207*(3), 158–163.

Bonawitz, E., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-stay, lose-sample: A simple sequential algorithm for approximating bayesian inference. *Cognitive psychology*, *74*, 35–65.

Budescu, D. V., & Rapoport, A. (1994). Subjective randomization in one- and two-person games. *Journal of Behavioral Decision Making*, *7*(4), 261–278.

Camerer, C. F. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton university press.

Cook, R., Bird, G., Lünser, G., Huck, S., & Heyes, C. (2012). Automatic imitation in a strategic context: players of rock–paper–scissors imitate opponents' gestures. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1729), 780–786.

Dyson, B. J. (2019). Behavioural isomorphism, cognitive economy and recursive thought in non-transitive game strategy. *Games*, *10*(3), 32.

Dyson, B. J., Wilbiks, J. M. P., Sandhu, R., Papanicolaou, G., & Lintag, J. (2016). Negative outcomes evoke cyclic irrational decisions in rock, paper, scissors. *Scientific Reports (Nature Publisher Group)*, *6*(1), 20479.

Frey, S., & Goldstone, R. L. (2013). Cyclic game dynamics driven by iterated reasoning. *PloS one*, *8*, e56416.

Hawkins, R. X. (2015). Conducting real-time multiplayer experiments on the web. *Behavior Research Methods*, *47*, 966-976.

Hawkins, R. X., Goodman, N. D., & Goldstone, R. L. (2019). The emergence of social norms and conventions. *Trends in cognitive sciences*, *23*, 158–169.

Morgenstern, O., & Neumann, J. V. (1953). *Theory of games and economic behavior*. Princeton university press.

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, *36*, 48–49.

Oey, L., Schachner, A., & Vul, E. (2019). Designing good deception: Recursive theory of mind in lying and lie detection. In *Proceedings of the 41st annual meeting of the cognitive science society*.

Smith, J. M. (1982). *Evolution and the theory of games*. Cambridge university press.

Stöttinger, E., Filipowicz, A., Danckert, J., & Anderson, B. (2014). The effects of prior learned strategies on updating an opponent's strategy in the rock, paper, scissors game. *Cognitive Science*, *38*(7), 1482–1492.

Wang, Z., Xu, B., & Zhou, H. J. (2014). Social cycling and conditional responses in the rock-paper-scissors game. *Scientific reports*, *4*, 5830.