# Extracting Cognitive Behavioral Features from Iterated Rock-Scissors-Paper Played Against An Artificial-Intelligence Agent

Kensuke Arai[1], Alik Widge[2], Suma Jacob[2] and Ali Yousefi[3]

***Abstract*— Quantification of human behavior in a social context may lead to discovery of subtle behavioral variations in a population that can be used to improve classification and screening for psychiatric disorders, and provide more accurate targeting in the development of interventions and biomedical treatments. However, it is difficult to study social interaction in a controlled, reproducible environment, as well as analyze the resulting behavior. In this research, we describe an experimental framework that utilizes a game of iterated Rock-Scissors-Paper played against an artificial intelligence agent, and a behavioral hypothesis of rule-switching to motivate analytical methods, that will extract behavioral features with potential clinical relevance from game data.**

***Clinical relevance*. Presently, psychiatric diagnosis involves subjective assessment of symptoms. Finding quantifiable, observable behavior that displays a spectrum in a population may be useful to differentiate and diagnose psychiatric illness.**

## I. INTRODUCTION

Psychiatric diagnosis is presently based on self-report and assessment by family members, teachers or psychiatrists [1], [2], [3]. However, biased response in self-report [4], low inter-rater reliability [5], [6] and inherent subjectivity, are known drawbacks to these assessments. There is recent efforts to instead observe behavior that can differentiate disease and healthy state that can be invoked in a controlled environment [7], [8]. Psychiatric illnesses are often most evident in social interaction, and thus social interactions are where it is most likely that differentiating behaviors can be found [9], [10]. However, studying social interaction is challenging for several reasons. First, describing a state of interaction requires finding a relevant and succinct set of variables that captures subtle social dynamic. Second, most behavioral tests involve measuring reactions to predetermined stimuli [10], [7]. This can be repeated across subject populations, but in a social interaction, there must be mutual interaction between environment and subject. Pairing subjects is a way to provide mutual interaction [11], but pairing cannot reproduce an environment across a subject population. To address these issues, we propose using AI-backed interactive, competitive games [12]. Competitive games can be described with a limited number of variables yet may still exercise some of the same social cognitive skills required in natural social interaction like reading anothers intentions [13], [14]. We argue replacing human opponents with an AI agent allows all subjects to play against the same opponent. Further, biomarkers can be derived from behaviors that can be isolated, by augmenting with bio-signals such as heart rate, respiration, EEG etc [15], [16], [17], [18].

In this research, we describe an experimental and analytical framework using an AI Rock-Scissor-Paper (RSP) agent to perform experiments that allows standardization of a social interaction across a subject population, and provide a model of behavior of how humans play the game of RSP in a repeated context against opponents that try to exploit them by finding patterns in their playing. Competitive games such as rock-scissors-paper have a game-theoretic optimal minimax solution, the Nash Mixed Equilibrium (NME) [19] that requires a player to choose actions randomly with equal probability. A building consensus is that humans are ill-equipped to behave randomly [20], [21], [9], and instead players adjust to their opponents. Players appear to be able to use historical information presented to them about novel opponents in best-of-3 scenario [11], though do not fully exploit regularities in their opponents over long, repeated games [25], yet monkeys have been observed to react to mixed NME computer opponents by selecting the same hand 90% of the time [26]. Humans may be relying on simple heuristic rules like Win-Stay-Lose-Switch, an approximate algorithm for sequential Bayesian inference [27]. The degree to which human players conform to simple heuristics like win-stay-lose-switch, have been found to be context dependent [28], [29], and the difference in relative flexibility of these two components suggest multiple decision making circuits are active in different contexts [30]. Researchers have considered evidence of heuristics beyond Win-Stay-Lose-Switch [31], and found evidence of individuality in the repertoire of heuristics used [32], [33], [34]. Building on the view of individuality in the heuristics players employ, we further hypothesize that humans adapt their use of heuristics during the course of multiple games [20] by switching to different heuristic rules. Rule-switching makes a testable prediction that competitive advantage is regained following rule switches, and we further find evidence that stable behavioral characteristic can be found from features derived from the time series of rule changes and other aspects of game play.

## II. METHODS

### A. Experiment

We used the online cognitive experimental platform test-mybrain.org [35] to recruit 1370 anonymous volunteers (human players, HP) from July to December 2021 to play 300-game rounds of rock-scissors-paper (rock beats scissors,

[1]Kensuke Arai and Ali Yousefi are with the Computer Science Department, Worcester Polytechnic Institute, Worcester, MA 01609, USA. Kensuke Arai is currently at Sumitomo Pharma, Osaka, Japan. kensuke.y.arai@gmail.com

[2]Alik Widge and Suma Jacob are with the University of Minnesota Medical School, Minneapolis, MN 55414, USA

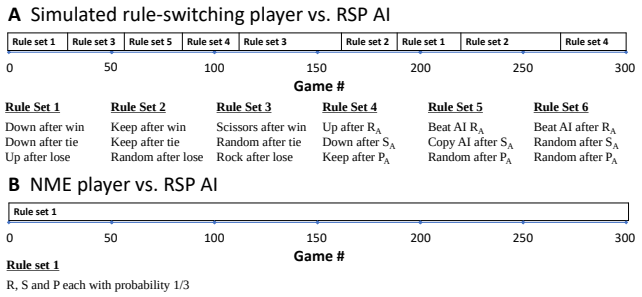


Fig. 1. Rule-switch schedule of 2 simulated HPs playing according to A) rule-switching and B) NME against the RSP AI. A) 6 rule sets, each containing complete instructions to generate next hand, switched episodically and randomly 8 times during course of 300-game round. Note Rule Sets 1 and 4 triggered by different events ($wtl$ vs. $r_A s_A p_A$ while Rule Sets 1 and 3 triggered by same $wtl$ events, but generate actions differently (relative to last hand DKU or concrete RSP). B) NME player plays randomly with equal probability through entire round.

scissors beats paper, paper beats rock) against an AI agent running on a web browser, playable on iOS, Android, Windows and Macintosh. The AI was comprised of 3-perceptrons each representing R, S and P [36], [37], whose internal parameters were updated online after each game, and that predicted the next HP move using the past 2 HP and AI moves. The games were self-paced, and data that suggested lack of engagement (average game-game interval $< 500$ms, more than 10 game-game intervals $> 1$ minute, or more than 15 consecutive repeated responses, were discarded. This left 192 game data where players completed a 300 game round and were deemed credibly engaged. The mean net wins over a round for HPs was -15.4, with standard deviation 28.1.

### B. Rule-switching hypothesis

Humans might approximate randomness by episodic switching between nearly deterministic rules of play [20] of the form "when $y$ happens, do X" (abbreviated as X$|y$) that pairs a triggering condition with an action. Given real-world cognitive limitations [27], [38], we further hypothesize that these action decisions are based on a limited representation of the most recent state of the game, for example the last outcome. To systematize this notion, we consider 3 classes of triggering condition: last outcomes (***w***in, ***t***ie, ***l***ose), HP's last hand (***r***ock, ***s***cissors, ***p***aper), or AI's last hand ($\boldsymbol{r}_{\text{A}}$ock, $\boldsymbol{s}_{\text{A}}$cissors, $\boldsymbol{p}_{\text{A}}$aper. We also consider 3 classes of HP actions: absolute (**R**ock, **S**cissors, **P**aper), relative to HPs last hand (**D**owngrade, ie after playing rock play scissors, **K**eep, ie after playing rock keep rock or **U**pgrade, ie after playing rock play paper), or relative to AI's last hand (**L**ose to, ie if AI played rock, play scissors next, **C**opy, ie mimic AI's last hand, or **B**eat, ie if AI played rock, play paper next). 9 rules are generated from a given trigger and action class pair (TACP), which we abbreviate like DKU$|wtl$. There are 3 sets of trigger and 3 sets of action classes, defining 9 TACPs, for a total of 9 TACPs $\times$ 9 rules/TACP $=$ 81 rules. 2 of the 3 TACPs sharing the same trigger class are re-descriptions of the same behavior, ie DKU$|rps$ and RPS$|rps$. They may represent different cognitive stances with different mental

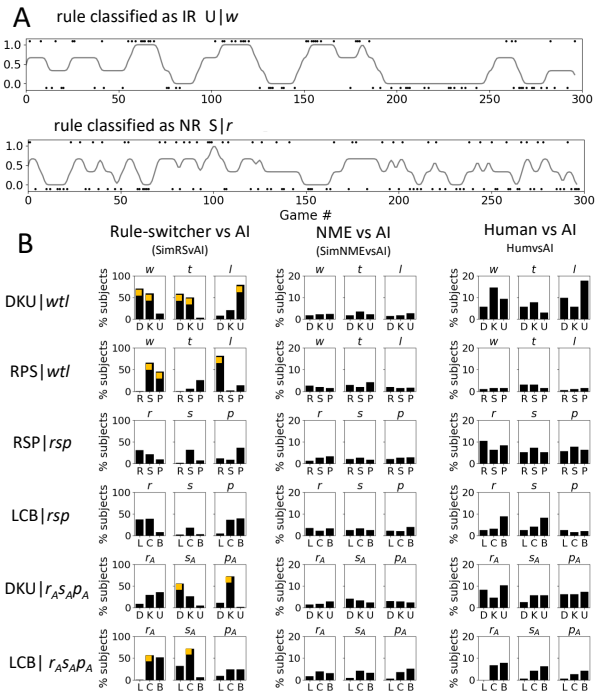


Fig. 2. A) Example traces of conditional probability for trigger-action rule detected to be IR and NR from a human subject. IR rules tend to have interspersed periods where probability is near 1 (switching), while NR rules rarely go near 1. B) shows the fraction in populations of rule-switcher simulations, NME simulations and human subjects in which each of the 6 x 9 rules were classified as IR. For the rule-switcher simulations, we have ground truth knowledge of what rules are actually in the repertoire (every simulation in population has identical repertoire), and these rules are marked with an orange box on the histogram. For the NME simulations, we expect near chance level detection equally across all rules. Presumably for human subjects, repertoire is heterogenous across population. Note the qualitative difference of the NME in comparison to the humans and rule-switchers.

economy necessary to generate them [32], but are behaviorally indistinguishable. Acknowledging this limitation in order to extract changes to observable behavior, we chose one TACP from the 2 isomorphic descriptions, resulting in the following 54 rules from 6 TACPs: DKU$|wtl$, RPS$|wtl$, RPS$|rps$, LCB$|rps$, DKU$|r_{\text{A}}s_{\text{A}}p_{\text{A}}$, and LCB$|r_{\text{A}}s_{\text{A}}p_{\text{A}}$ from here on. The well-known "win-stay-lose-switch" heuristic, an approximate algorithm for online Bayesian inference [27], takes the rules $K|w$ and $D$ or $U|l$ from DKU$|wtl$. How strongly a player conforms to these rules on average has been used to characterize changes in competitive behavior in various conditions [39], [30], [29], [28] and to argue for multiple drivers of decision making in the brain. LCB$|r_{\text{A}}s_{\text{A}}p_{\text{A}}$ are known as the Cournot Best, Second Best and Worst Responses. We hypothesize players may be using several of these rules in short bursts. A player may play 20 games always following a win with a hand that beats their last hand (U$|w$), then for the next 15 games, switch to keeping the same hand following a win (K$|w$). To our knowledge, consideration of these rules being used dynamically as part of a larger repertoire and attempting to uncover the hidden changes to behavior, is a novel research direction.

### C. Simulated player vs AI, Human players vs AI

To test our ability to detect rules present in game data, in addition to collecting data from human players vs AI, we also simulated two idealized types of human players against our AI. We simulated 1) a rule-switcher that can choose and episodically switches rules from a rule repertoire as illustrated in Fig. IA, and 2) an NME player that plays RSP with equal probability randomly throughout the round, Fig. IB. We refer to the data collected in these experiments as HPvAI, SimRSvAI and SimNMEvAI, respectively. In SimRSvAI, we know which of the 54 possible rules are in the simulated HPs repertoire, and we know the exact moment rule sets were switched. In the SimNMEvAI, we know there is no conditional structure in the data, as well as no rule changes.

### D. Inferring rule repertoire

We hypothesize that player tendencies are reflected in the types of rules they utilize [39], and that it is nearly deterministic for several hands during its use, before a switch to a different rule is made. We expect that a player uses only a small subset of the 54 possible rules. We call this subset the rule repertoire. If these assumptions hold, it is possible to detect with reasonable fidelity whether a rule is used at all during the round, and if so, when it is in use. For concreteness of discussion, we consider the rule "Upgrade after winning" (U$|w$), and estimate the time evolution of $p$(U$|w$) during a 300-game round. We first make a temporally ordered list of all $w$s (not necessarily at consecutive game #s). We estimated $p$(U$|w$) as a function of game # using a sliding window of size $B$. We look at the first $B$ $w$s in the list, count the number of U actions that follow each one, and divided by B to obtain a probability, and assign this probability to $p$(U$|w$)(game # of middle $w$ of window), and slide by 1 in the list of $w$s and repeat the process for $w$ #2, #3, . . . #$B$+1, et cetera. Because consecutive $w$ conditions are not necessarily occuring at consecutive games, the $p$(U$|w$) is defined at intermittent games. We interpolate between the points at which they are defined, and smooth in anticipation of taking time derivatives later. Because we do not expect rules to be kept for too long, we chose $B = 3$ in our analysis. Varying $B$ between 3 and 5 does not qualitatively change the results of our analysis.

If one of the rules employed by the player in a round is U$|w$, we expect to see an excess of consecutive U actions following triggering $w$s. Shuffling the game order should destroy the trigger-action relationship, resulting in a reduction of consecutive U actions following triggering $w$s in shuffled data. We test for the excess by counting the total duration where $p$(U$|w$)(game #) $> p_{th}$, and compare this to the bootstrap distribution created from the shuffled data. Each of the 6 x 9 rules were tested, and we classify a rule whose excess of consecutive U actions following a triggering $w$ is significant, as in-repertoire (IR). This rule is considered to be part of the repertoire of rules HP uses, and the remaining rules we refer to as non-repertoire (NR). For our analysis, we chose $p_{th}$ = 0.6. Fig. 2A top and bottom trace show conditional probabilities of rules from an HP that were classified as IR and NR, respectively.

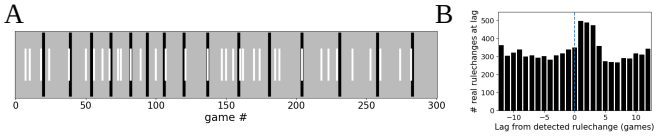


Fig. 3. A) Detected (white) and actual (black) rule change times shown for an example round from SimRSvAI. B) Histogram of ground truth rule change times triggered by detected rule change times, showing peak near detected rule change time. The current rule change detection method produces many false positive events, though most black ticks also have a white tick nearby, indicating that real rule changes are rarely missed.

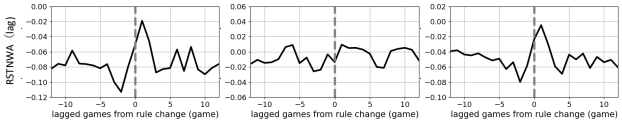


Fig. 4. RSTNWA for SimRSvAI, SimNMEvAI and HPvAI. Prominent increase in wins occurs near instants of detected rule switches in SimRSvAI and HPvAI but not in SimNMEvAI, as expected. Because RSTNWA is looking at the change in net wins, we excluded detected switches of rules from the TACPs DKU$|wtl$ and RSP$|wtl$ from our triggers, because any effect on the net win rate near those rule switches might be due to the fact that the triggers of the rules themselves are wins and losses. Note the differences in offset in the 3 RSTNWA. RSTNWA offset for SimNMEvAI is near 0 because the NME strategy is the minimax solution that can neither exploit not be exploited. RSTNWA offset of both SimRSvAI and HPvAI are $< 0$, indicating that on average, AI is exploiting both the simulated rule switcher and HPs.

### E. Detecting Rule Switches

After we find which rules are IR, we next detect at which games switching in these IR rules may have occurred. In Fig. 2A top trace, rapid swings from nearly 0 to 1 or 1 to 0 around games 50, 75, 100, 125, 150, 190, 250, 260 are what we envision rule changes would look like. A simple criteria to detect when rule switching occurred is to find the local maxima of the absolute value of the time derivatives for each IR rule. We consider the collection of these rule-change time points are previously undetected changes in behavior that underlies the mechanism employed by humans to appear random against a competitor.

## III. Results

### A. Rule repertoire detection

Fig. 2B left, middle and right figure compares the fraction of all the data in SimRSvAI, SimNMEvAI, HPvAI, respectively, in which each of the 54 rules was classified as IR. There is a noticeable difference in the number of IR rules detected in the 3 types of data, with the fewest IR rules detected in SimNMEvAI, as expected. The identities of the detected IR rules are also found to be reasonably close to what is actually in the repertoire for SimRSvAI. Also notable is that the probability of any given rule detected as IR is much more uniform in SimRSvAI, as expected. HPs, unlike the population of simulated rule-switchers that all had identical rule repertoires, presumably are not identical,

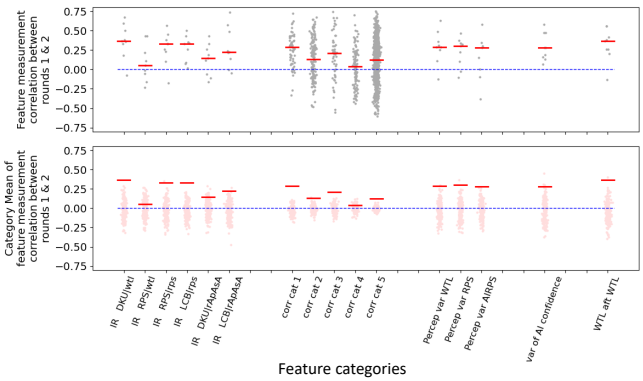


Fig. 5. Correlation between features in repeated measurements. Features systematically generated and organized into categories measured from 18 subjects who played 2 rounds of RSPvAI. Top plot shows the correlation between the 1st and 2nd measurement for features in each category. Red line is the mean correlation for all features in the category. Bottom plot shows the same mean correlation for different shuffles of the subject index of the 2nd measurement, while red is the mean for unshuffled (same as top plot). Blue line in both top and bottom is 0.

and hence the fraction of any rule being IR is much lower in HPvAI than SimRSvAI. It is notable that rules of type RPS$|wtl$ are much less likely to be IR in HPvAI, suggesting that our analysis is picking up a general feature of human competitive behavior.

### B. Rule change detection

Once the IR rules are detected, known changes in the rule-switcher simulations were assessed by comparing them to the detected rule-switches, Fig. 3. The large number of false positive detections is to be expected, given the simple detection methodology presented in this paper. A latent state model [15], [40] would improve upon the current sensitive but non-specific algorithm, and our current exploratory analysis will motivate more sophisticated future modeling and guide the design of the latent state.

### C. Advantage regained at rule change

An expected consequence of episodic rule changing is that advantage is temporarily regained immediately after a rule change. Because the AI is constantly updating its internal model of the human HP, sudden changes in human behavior should result in the internal model of the AI becoming obsolete. We define the rule-switch triggered net win average (RSTNWA), constructed by stacking the segments of the timeseries of outcomes preceding and following detected rule switches. At each lag from the triggering rule switch, the number of losses is subtracted from the number of wins at that lag, the netwins, is divided by the combined number of wins and losses, RSTNWA(lag) $=$ #wins at lag $-$ #losses at lag)/(#wins at lag $+$ #losses at lag). Fig. 4A right shows RSTNWA abruptly increasing at moment of rule change, as expected. Fig. 4B right shows RSTNWA for the random player, showing no increase at rule change time, while Fig. 4C shows qualitative similarity as Fig. 4A.

### D. Features derivable from RPSvAI are characteristic of the subjects

We aim to extract behavioral features that are characteristic of the subject from RPSvAI game data. The features we considered are 1) which rules are in a subjects repertoire, 2) how a subject coordinates the switches of multiple rules over the course of a round, 3) how the AIs confidence in a prediction fluctuates over time, and 4) the conditional outcome probabilities, Fig. 5. From the 54 rules, a total of 1485 features were systematically created. We created 2 types of features looking at the standard deviation of the conditional probabilities (6 TACPs x 9 rules = 54 features) and the correlation coefficient between pairs of rule conditional probabilities (54 x 53 / 2 features). The correlations were grouped by 1) terms from same TACP, same condition (ie $p(\mathrm{D}|w)$ and $p(\mathrm{U}|w)$), 2) terms from same TACP, different condition, (ie $p(\mathrm{U}|w)$ and $p(\mathrm{U}|t)$), 3) terms from different TACP sharing condition class, same condition, (ie $p(\mathrm{U}|w)$ and $p(\mathrm{R}|w)$) 4) terms from different TACP sharing condition class, different condition, (ie $p(\mathrm{U}|w)$ and $p(\mathrm{R}|t)$) and 5) terms from different TACP that do not share condition class (ie $p(\mathrm{U}|w)$ and $p(\mathrm{R}|t)$). Further, we looked at the decision confidence of the AI perceptrons following $wtl$, $rsp$, $r_{\mathrm{A}}s_{\mathrm{A}}p_{\mathrm{A}}$ (3 x 3 x 3 = 27 features), and the conditional outcome (3 x 3 = 9 features) for a total of 1521 features. We did not account for whether a rule was IR or NR in each subject, as this would make it difficult to compare the same set of features across all 18 subjects, and with better methods to detect rules, we do not expect the conditional probability to be the best way to characterize rules. However for this analysis, we deemed it sufficient. We measured these features in 2 separate rounds, and measured the correlation in these 2 measurements. We expect repeated measurement of a feature to be relatively stable, so we expect the correlations to be positive. Because only 18 of 192 subjects played more than 1 round of RSPvAI, our analysis is preliminary, but we found consistent trend towards positive correlation between measurements made in 2 separate rounds, suggesting we are capturing behavioral characteristics of individuals, Fig. 5, top figure.. The average in 1521 total features of the correlation coefficient between 1st and 2nd measurements was 0.2. For each category of feature, While this is not a high correlation, the results are not discouraging for several reasons. First, we constructed all features systematically, and it is likely that not all features constructed in such fashion is a characteristic, stable behavioral feature. Second, it may be a possibility that 300 game rounds is not long enough to accurately measure some features, especially how certain rules are coordinated with others. Most players were able to complete a round in under 7 minutes, and longer rounds or possibly multiple rounds are a possibility.. Also, better methodology used to find IR rules and detect rule changes, is an obvious area where improvements are possible that may lead to better quality in the extracted features.

## IV. CONCLUSION

We developed an experimental framework using a competitive game played against an AI agent to incorporate social interaction in a repeatable and controlled way. We also put forth a model of behavior, rule-switching, that builds testable hypothesis about the deviations from the Nash Mixed Equilibrium strategy that allows us to interpret, quantify and extract behavioral features that potentially characterizes individual subjects. We believe that our analysis methods, while unsophisticated, explicitly exposes a randomization mechanism implicitly assumed, for playing against an opponent that can exploit regularities, and lays the groundwork for future, more sophisticated statistical latent modeling of behavior. We have also shown that, in addition to features derived from human behaviors, those derived from the internal AI parameters shaped by game play against the HP, may also characterize each individual. Altering the AIs exploitability may reveal differences in how advantageous rules switched to are, and offer a more fine-grained metric of competitive strength than simple net wins that may provide further quantitative features of behavior.

## ACKNOWLEDGMENT

We thank Shigeru Shinomoto for assistance in development of our online experiment, and Uri Eden for discussions about the analysis of data.

## REFERENCES

[1] Stanley R Kay, Abraham Flszbein, and Lewis A Opler. The positive and negative syndrome scale (panss) for schizophrenia. *Schizophrenia Bulletin*, 13, 1987.

[2] Simon Baron-Cohen, Sally Wheelwright, Richard Skinner, Joanne Martin, and Emma Clubley. The autism-spectrum quotient (aq): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31, 2001.

[3] Rosa A. Hoekstra, Anna A.E. Vinkhuyzen, Sally Wheelwright, Meike Bartels, Dorret I. Boomsma, Simon Baron-Cohen, Danielle Posthuma, and Sophie Van Der Sluis. The construction and validation of an abridged version of the autism-spectrum quotient (aq-short). *Journal of Autism and Developmental Disorders*, 41:589–596, 2011.

[4] Margot M. Williams, Richard Rogers, Allyson J. Sharf, and Colin A. Ross. Faking good: An investigation of social desirability and defensiveness in an inpatient sample with personality disorder traits. *Journal of Personality Assessment*, 101:253–263, 2019.

[5] Ross M G Norman, Ashok K Malla, Leonard Cortese, and Fernando Diaz. A study of the interrelationship between and comparative interrater reliability of the saps, sans and panss. *Schizophrenia Research*, 19:73–85, 1996.

[6] Paul S.S. Russell, Anna Daniel, Sushila Russell, Priya Mammen, Julie S. Abel, Lydia E. Raj, Satya Raj Shankar, and Naveen Thomas. Diagnostic accuracy, reliability and validity of childhood autism rating scale in india. *World Journal of Pediatrics*, 6:141–147, 5 2010.

[7] Yasuki Ono, Kiwamu Kudoh, Takashi Ikeda, Tetsuya Takahashi, Yuko Yoshimura, Yoshio Minabe, and Mitsuru Kikuchi. Auditory steady-state response at 20 hz and 40 hz in young typically developing children and children with autism spectrum disorder. 2020.

[8] A. David Redish, Adam Kepecs, Lisa M. Anderson, Olivia L. Calvin, Nicola M. Grissom, Ann F. Haynos, Sarah R. Heilbronner, Alexander B. Herman, Suma Jacob, Sisi Ma, Iris Vilares, Sophia Vinogradov, Cody J. Walters, Alik S. Widge, Jennifer L. Zick, and Anna Zilverstand. Computational validity: using computation to translate behaviours across species. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 377:20200525, 2 2022.

[9] Kwangyeol Baek, Yang Tae Kim, Minsung Kim, Yohan Choi, Minhong Lee, Khangjune Lee, Sangjoon Hahn, and Jaeseung Jeong. Response randomization of one- and two-person rock-paper-scissors games in individuals with schizophrenia. *Psychiatry Research*, 207:158–163, 5 2013.

[10] Alexis B. Craig, Emily Grossman, and Jeffrey L. Krichmar. Investigation of autistic traits through strategic decision-making in games with adaptive agents. *Scientific Reports*, 7, 12 2017.

[11] Dimitris Batzilis, Sonia Jaffe, Steven Levitt, John A. List, and Jeffrey Picel. Behavior in strategic settings: Evidence from a million rock-paper-scissors games. *Games*, 10:1–34, 2019.

[12] Lei Wang, Wenbin Huang, Yuanpeng Li, Julian Evans, and Sailing He. Multi-ai competing and winning against humans in iterated rock-paper-scissors game. *Scientific Reports*, 10:1–8, 2020.

[13] David Premack and Guy Woodruff. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 4:515–526, 1978.

[14] Wako Yoshida, Ray J Dolan, and Karl J Friston. Game theory of mind. *PLoS Computational Biology*, 4:1–14, 2008.

[15] Matthew S D Kerr, Sandya Subramanian, Zachary Fitzgerald, Kevin Kahn, Matthew A Johnson, Pierre Sacr, Ernst Niebur, Uri T Eden, Jorge A Gonz, John T Gale, and Sridevi V Sarma. Risk-taking bias in human decision-making is encoded via a right left brain push pull system. 116, 2019.

[16] Hiroki Fukui, Toshiya Murai, Jun Shinozaki, Toshihiko Aso, Hidenao Fukuyama, Takuji Hayashi, and Takashi Hanakawa. The neural basis of social tactics: An fmri study. *NeuroImage*, 32:913–920, 2006.

[17] Shunsuke Sugiyama, Kazutaka Ohi, Ayumi Kuramitsu, Kentaro Takai, Yukimasa Muto, Tomoya Taniguchi, Tomoaki Kinukawa, Nobuyuki Takeuchi, Eishi Motomura, Makoto Nishihara, Toshiki Shioiri, and Koji Inui. The auditory steady-state response: Electrophysiological index for sensory processing dysfunction in psychiatric disorders. *Frontiers in Psychiatry*, 12, 3 2021.

[18] Hongli Wang, Heather K. Ortega, Huriye Atilgan, Cayla E. Murphy, and Alex C. Kwan. Pupil correlates of decision variables in mice playing a competitive mixed-strategy game. *bioRxiv*, 3:6, 2021.

[19] Federico Sanabria and Eric Thrailkill. Pigeons (columba livia) approach nash equilibrium in experimental matching pennies competitions. *Journal of the Experimental Analysis of Behavior*, 91:169–183, 2009.

[20] Robert L. West, Christian Lebiere, and Dan J. Bothell. Cognitive architectures, game playing, and human evolution, 2005.

[21] Elham Jokar and Mohammad Mikaili. Assessment of human random number generation for biometric verification. *Journal of Medical Signals and Sensors*, 2:82–87, 2012.

[22] Bin Xu, Hai jun Zhou, and Zhijian Wang. Cycle frequency in standard rock paper scissors games : Evidence from experimental economics. *Physica A*, 392:4997–5005, 2013.

[23] Seth Frey and Robert L. Goldstone. Cyclic game dynamics driven by iterated reasoning. *PLoS ONE*, 8, 2 2013.

[24] Zhijian Wang, Bin Xu, and Hai Jun Zhou. Social cycling and conditional responses in the rock-paper-scissors game. *Scientific Reports*, 4, 7 2014.

[25] Erik Brockbank and Ed Vul. Humans fail to outwit adaptive rock, paper, scissors opponents. page 43, 2021.

[26] Daeyeol Lee, Benjamin P Mcgreevy, and Dominic J Barraclough. Learning and decision making in monkeys during a rock paper scissors game. 25:416–430, 2005.

[27] Elizabeth Bonawitz, Stephanie Denison, Alison Gopnik, and Thomas L. Griffiths. Win-stay, lose-sample: A simple sequential algorithm for approximating bayesian inference. *Cognitive Psychology*, 74:35–65, 2014.

[28] Benjamin James Dyson, Jonathan Michael, Paul Wilbiks, Raj Sandhu, Georgios Papanicolaou, and Jaimie Lintag. Negative outcomes evoke cyclic irrational decisions in rock, paper, scissors. *Nature Publishing Group*, pages 1–6, 2016.

[29] Vincent Srihaput, Kaylee Craplewe, and Benjamin James Dyson. Switching competitors reduces win-stay but not lose-shift behaviour: The role of outcome-action association strength on reinforcement learning. *Games*, 11:1–10, 2020.

[30] Lewis Forder and Benjamin James Dyson. Behavioural and neural modulation of win-stay but not lose-shift strategies as a function of outcome value in rock, paper, scissors. *Scientific Reports*, 6, 9 2016.

[31] Erik Brockbank and Edward Vul. Recursive adversarial reasoning in the rock, paper, scissors game, 2020.

[32] Benjamin J. Dyson. Behavioural isomorphism, cognitive economy and recursive thought in non-transitive game strategy. *Games*, 10, 9 2019.

[33] Hanshu Zhang, Frederic Moisan, and Cleotilde Gonzalez. Paper-rock-scissors: an exploration of the dynamics of players strategies. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 64:268–272, 2020.

[34] Julia Poncela-Casasnovas, Mario Gutirrez-Roig, Carlos Gracia-Lzaro, Julian Vicens, Jess Gmez-Gardees, Josep Perell, Yamir Moreno, Jordi Duch, and Angel Snchez. Humans display a reduced set of consistent behavioral phenotypes in dyadic games. *Science Advances*, 2:1–9, 2016.

[35] Laura Germine, Ken Nakayama, Bradley C. Duchaine, Christopher F. Chabris, Garga Chatterjee, and Jeremy B. Wilmer. Is the web as good as the lab? comparable performance from web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin and Review*, 19:847–857, 10 2012.

[36] Shigeru Shinomoto. *Theory of Data Analysis, from Prediction to Simulation (Japanese)*. Iwanami Shoten, 2002.

[37] Matthew F Rutledge-Taylor and Robert L West. Using dshm to model paper, rock, scissors. pages 2341–2346, 2011.

[38] Jori Barash, Isabelle Brocas, Juan D. Carrillo, and Niree Kodaverdian. Heuristic to bayesian: The evolution of reasoning from childhood to adulthood. *Journal of Economic Behavior and Organization*, 159:305–322, 2019.

[39] Hanshu Zhang, Frederic Moisan, and Cleotilde Gonzalez. Rock-paper-scissors play: Beyond the win-stay/lose-change strategy. *Games*, 12, 9 2021.

[40] Kensuke Arai and Robert E. Kass. Inferring oscillatory modulation in neural spike trains. *PLoS Computational Biology*, 13:1–31, 2017.