Applying Occam's razor to paper (and rock and scissors, too): Why simpler models are sometimes better

Antonio Napoli (antonio.napoli@phd.units.it)
Danilo Fum (fum@units.it)
Dipartimento di Psicologia, Università degli Studi di Trieste
via S. Anastasio 12, I-34134 Trieste, Italy

Abstract

A commonly held idea is that people engaged in guessing tasks try to detect sequential dependencies between the occurring events and behave accordingly. For instance, previous accounts of the popular Rock Paper Scissors game assume that people try to anticipate the move an opponent is likely to make and play a move capable of beating it. In the paper we propose that players modulate their behavior by reacting to the effects it produces on the environment, i.e., that they behave exactly as they do in non competitive situations. We present an experiment in which participants play against a computer controlled by different algorithms and develop a procedural model, based on the new ACT-R utility learning mechanism, that is able to replicate the participants' behavior in all the experimental conditions.

Keywords: Rock-Paper-Scissors, reinforcement learning, procedural learning, ACT-R, neural nets.

Introduction

The capability of adapting to changes occurring in the environment and to anticipate future events constitutes a critical factor for organisms' survival, and humans and animals have been tuned by natural selection to become receptive to subtle variations in the external contingencies. Adaptivity and proactivity are realized essentially through a process of selection by consequences—named also law of effect (Thorndike, 1898), operant conditioning (Skinner, 1938) or reinforcement learning (Sutton & Bartho, 1998)—i.e., on the idea that organisms modulate their behavior by reacting to the effects it produces on the environment.

Some predictions organisms routinely make concern the behavior of other organisms. A particular situation in which such predictions are useful is given by competitive games. It's obvious that, if we knew in advance the move our opponent is going to make, our life would become easier. In the paper we deal with Rock Paper Scissors (aka Roshambo), a competitive game that, while being extremely simple to describe and play, presents a series of interesting features when considered from a cognitive point of view.

The following section presents the essentials of the game and describes some strategies that have been suggested to play it effectively. Next, we review previous studies which investigated the behavior of human players in this task and proposed some models to explain it. As it will become apparent in the following, a common theme underlying this work is that people attempt to succeed at the game by trying to anticipate the move the opponent is likely to make and playing a move capable of beating it. We advance, on the other hand, a simpler explanation for the players' performance which relies on the same principle of selection by consequences that explains most of the behavior in non competitive situations. We present an experiment in which participants play against a computer controlled by different algorithms and develop a procedural model based on the new ACT-R utility learning mechanism that is able to replicate the participants' behavior in all the experimental conditions.

Rock Paper Scissors

Rock Paper Scissors (henceforth RPS) is a competitive two-person game which is played through a series of turns in which players make their moves simultaneously. The outcome of each turn is determined as follows: Rock beats Scissors, Scissors beats Paper, but Paper beats Rock. If both players make the same move, the turn is considered as a tie. That's all, as far as the game's rules are concerned.

In RPS no move—no "pure strategy", in terms of Game Theory (Von Neumann & Morgenstern, 1944)—can be considered as the best to play. Concepts like "better", "bigger", "stronger" and similar are possible only referring to sets for which a partial ordering could be established, and this requires the existence of a transitive relation among set members, an eventuality that cannot be realized in RPS where the relation "beats" originates a closed loop.

Considered from the point of view of the Game Theory, RPS is classified as a two-person zero-sum game. For all games of this kind there exists a solution, i.e., a rule or norm that prescribes how the game should be played. Assuming perfectly rational players, the solution coincides with the Nash equilibrium at which neither player could hope to achieve a better performance by modifying their strategy, In case of RPS, the Nash equilibrium is reached by choosing the three possible moves randomly with equal probability, i.e., by playing a mixed strategy through a stochastic combination of the pure strategies.

While game theorists could consider RPS as a trivial game, there are two facts that make it intriguing from a cognitive point of view. First of all, humans are notoriously bad at generating random moves (Rapoport & Budescu, 1997; Wagenaar, 1972), so theorists could not easily practice what they preach. Being unable to play randomly, humans necessarily display sequential dependencies among the moves they make that could be exploited by a clever opponent. Second, the mixed strategy has the advantage that no strategy can beat it but it also has the disadvantage that there is no strategy that it can beat. In other words, it guarantees a break-even result in the long run, regardless of how strong (or how weak) the opponent is, but it does not allow a player to reach consistent wins.

In fact, aficionados consider RPS a game of wit, not a game of chance. Even a cursory look at the web site of the World RPS Society (www.worldrps.com) or a quick skim of *The official rock paper scissors strategy guide (*Walker & Walker, 2004) should convince that RPS experts use their insight to try to anticipate the opponent's move, possibly recurring to particular sequences of moves to try to induce predictable responses in the other player. The problem is that, to exploit a weakness in the opponent's play, you need to make non-random moves, which makes you vulnerable.

A clearer idea about which strategies could succeed at the game may be obtained by looking at the results of the First and Second International RoShamBo Programming Competition—held at the University of Alberta, Canada, in September 1999 and July 2000, respectively—two tournaments between computer programs playing RPS in which each program competed against all others. Because organizers enrolled in the competition some really weak programs that produced easily predictable move sequences, a program that played the optimal strategy without trying to exploit the competitors' deficiencies (running at the same time the risk to expose its owns) could reach only weak results. It should be noted that all programs could store the complete sequence of moves played by themselves and by the opponent, a feature which human players, due to their memory limitations, cannot easily rely upon.

The programs adopted essentially two high-level strategies to choose their moves. The first one was based on pattern-matching and tried to exploit the statistical regularities occurring in the sequence of moves produced by the opponent. The second one relied on some kind of meta-reasoning to determine how the opponent would choose its move. One of the most complicated strategies of this kind was represented by the so called Sicilian-reasoning according to which a program tried to figure out the competitor's move by assuming that it will think like itself, taking however than into account the fact that the competitor was likely to use Sicilian reasoning too, and giving thus raise to a “I know that you know that I know ...” recursive pattern. This approach was very effective and programs adopting it ranked among the best.

While computer programs could shed light on how RPS should be played by perfectly rational agents with unlimited memory, we could ask how individual with bounded rationality, cognitive limits and emotions (i.e., normal people) really play the game.

Previous work

In the last decade Robert West, with Christian Lebiere and coworkers, produced a series of studies (West, 1999; Lebiere & West, 1999; West & Lebiere, 2001; Routledge-Taylor & West, 2004, 2005; West, Stewart, Lebiere & Chandrasekharan 2005) focused on the analysis of human behavior in the RPS and on the attempts to simulate it. These studies present several experiments whose results are explained through models that differ slightly from paper to paper. Through their comparative exam it is possible, however, to extract a unitary view and a coherent story that we are now going to tell.

According to the authors, people engaged in the RPS, and similar guessing tasks, try to detect regularities in the occurring events—in our case in the sequence of moves made by the opponent—and use this information to modulate their behavior. If both players use the same strategy of sequence detection, they enter in a state of reciprocal causation in which each player tries to influence the opponent's behavior while being, simultaneously influenced by it. The result is a dynamic, coupled system capable of generating patterns of interaction that could not be explained by looking at each system in isolation.

The players' behavior could be explained and replicated by a model capable of storing a variable number of previous opponent's moves. Differently from the computer programs playing the same game, the model has a reduced memory buffer whose capacity constitutes a critical factor in determining its behavior. A model which stores only the previous opponent's move is said to be a Lag1 model, if it stores the previous two moves is said to be Lag2, and so on.

The intuitive idea behind the models is that, if players could figure out what an opponent, having made the moves represented in the memory buffer, is going to do, they should make the move capable of beating it. This idea has been realized and implemented in different ways.

West (1999), Lebiere & West (2001) and Rutledge-Taylor & West (2004) used a two layers neural net which received in input the opponent’s moves and gave as output the move made by the player. The input layer comprised a number of node triples (each node representing Rock, Paper or Scissor, respectively) corresponding to the number of opponent's moves the model could store: one three-nodes group for Lag1—storing only the last move—two groups for Lag2— storing the last and the last but one moves—etc. Each input node could have a value of 0 or 1. More particularly, for each input triple, the node corresponding to the move made by the opponent received an activation value of 1 while the remaining two nodes got a 0.

All the nodes of the input layer were linked to the three nodes, one for each move, constituting the output layer. The weights of the links connecting the input and output nodes were initialized to 0 (in West, 1999 and West & Lebiere 2001) or were assigned a value randomly chosen from the set {-1, 0, +1} (in Rutledge-West, 2004). The value of an output node was determined by summing the weights of the links coming from the activated input nodes, i.e., from those input nodes having their activation set to 1. The network returned the move associated with the highest-value output node, possibly making a random choice in case of multiple nodes with the same activation.

After each choice, the link weights were adjusted according to the outcome. Two main policies were followed for updating the links. In the "passive" models, wins were rewarded by adding 1 to the weights of the links coming from the activated nodes (i.e., input nodes with a value of 1) and leading to the output node corresponding to the chosen move, losses were punished by subtracting 1, while ties kept all the links unvaried. "Aggressive" models, on the other hand, treated ties like losses and subtracted 1 to the links connecting the activated input nodes with the non-winning output node.

West & Lebiere (2001) carried out a series of experiments in which human participants played against different versions of the model and compared these results with those obtained by having different models compete against each other. In general, games in which identical versions of the models were pitted against each other ended in a tie. On the other hand, it was found that a broader memory span or a more aggressive attitude provided a competitive advantage: Lag2 models were able to reliably defeat Lag1 models while aggressive versions were superior to passive ones. Interestingly, the advantage provided by an extra lag or a more aggressive attitude were additive and approximately equal in magnitude.

Coming to human players, they were able to win on the average 9.99 turns (after a 300 turns game) more than the Aggressive Lag1 and 11.14 turns (after 287) more than the Passive Lag2 when pitted against these algorithms. According to West & Lebiere (2001), humans perform like the Aggressive Lag2 that constitutes, according to the authors, the best model for their behavior. Humans showed indeed a small but statistically significant trend to lose, instead of tie, against this model, but this effect was attributed to the fact that they were not able, due to lack of motivation and/or fatigue, to play in the same consistent manner as their computerized opponent.

These findings were congruent with those reported in those papers (West, 1999, Lebiere & West, 1999; Rutledge-Taylor & West, 2005, West, Stewart, Lebiere, & Chandrasekharan, 2005) that utilized models based on the ACT-R (Anderson & Lebiere, 1998) cognitive architecture. The idea that RPS is played exploiting the last moves to try to anticipate the next one to is maintained in these models, but in this case the sequence of moves is stored and retrieved not through a neural net but by taking advantage of the ACT-R declarative mechanism, while the model's choices, instead of being driven by the nodes' activation values, are demanded to the ACT-R procedural system.

The ACT-R declarative memory stores chunks containing the opponent's previous patterns and a prediction for the next move. The most important procedure used by the models is the following (slightly adapted from Lebiere & West, 1999, p. 297, and of intuitive significance):

Sequence Prediction

IF no move has been played
and the opponent last played move(s) L2 (and L1)
and move(s) L2 (and L1) are followed by move L
and move L is beaten by move M
THEN play move M.

For each turn, the model recalls the most active chunk matching the two (for Lag2 models) or the last (for Lag1 models) opponent's move(s). The model notes the move predicted by the chunk and plays the move that beats it. After both players have made their choice, a new chunk containing the update moves and the corresponding prediction is formed, or a previously existing chunk already storing the same information is reinforced.

Using this approach, Lebiere & West (1999) were capable of replicating the results of West (1999) while Rutledge-Taylor & West (2005) constructed several models capable of replicating the results of West & Lebiere (2001).

The Experiment

Our interest for RPS arose after a series of experiments (e.g., Fum & Stocco, 2003; Fum, Napoli, & Stocco, 2008,; Stocco, Fum, & Napoli, 2009) which found the participants' behavior heavily influenced by the principle of selection by consequences. These experiments, however, dealt with non competitive situations, i.e., situations that could be classified as "a *one-person game*, sometimes called a *game against natur*e, wherein a single player makes a decision in the face of an environment assumed to be indifferent or neutral" (Zagare, 1984, p. 11). To investigate whether the same principle could hold in competitive situations like RPS, we established the following experiment.

Participants played three rounds of RPS against a computer controlled in each round by a different program. A first group of participants (in the "Classic" condition) interacted with the computer through an interface which adopted the usual symbols, i.e. clinched fist for Rock, flat hand for Paper, and closed hand with extended index and Scissors. Participants were instructed about the rules of the game and were told that the computer was following in each round a different strategy that could however be defeated, at least in some cases. Immediately after the participants' made their choice, the computer move was displayed together with the turn outcome. Wins allowed the participants to gain one point (+1), losses were punished by the same amount (-1), while ties left the score unaltered (0).

A second group of participants was engaged in the same task arranged, however, as a game against nature. In this "Nature" condition the computer was presented not as a competitor but as a neutral game device allowing participants gain or lose points. In fact the programs the computer used were exactly the same of the previous condition. Instead of the classic RPS symbols, however, participants saw on the screen three geometric figures representing a sphere, a cube and a pyramid. At the beginning of each round the computer randomly matched each figure with an RPS move and behaved accordingly. Participants were told that they could obtain in each turn a score of +1, 0 or -1. It was also said that the criterion according to which the computer assigned scores to figures could not be easily guessed and that, in any case, it would change in each round. By relying on their "intuition" participants had to try to obtain in each round as many points as possible.

Another difference between these conditions, in addition to the setting and the use of different move images, was given by the fact that the computer, instead of the move it made in each turn, displayed the complete payoff matrix allowing participants to see both the score gained by their move and the scores they could have gained by making the alternative choices. Suppose a participant chose the sphere (matched for that round with Scissors) and the computer the cube (matched with Paper, while the pyramid was matched with Rock). The outcome was shown by displaying +1 in correspondence to the sphere (because Scissors beat Paper), 0 near the cube (each move ties with itself) and -1 near the pyramid (because Rock beats Scissors).

In the experiment a third condition (named "Implicit") was utilized that was presented against as a competitive situation in which participants had to choose one of the new symbols (the sphere, the cube and the pyramid) displayed on the screen. Participants were told that each figure could beat another figure, tie with itself, and lose against the third one but the payoff matrix was not revealed and had to be discovered by playing the game.

In summary, participants played against the computer, controlled by the same algorithms, in a classic RPS game, in a situation disguised as a game against nature, and in a competitive framework with unknown payoffs. We wanted to establish how participants would perform in the three conditions and, in particular, whether their behavior should be explained by using models of a different kind.

Method

Participants and design. Sixty students (37 males) enrolled at the University of Trieste, Italy, were recruited as participants. Their age varied between 18 and 32 years (*M*=21.4, *SD*=3.7). Participants were randomly assigned to one of the three experimental conditions (Classic, Implicit, and Nature) in which they were engaged in three RPS rounds, each one against a different algorithm whose order was counterbalanced between rounds. The experiment followed therefore a 3x3 mixed design with Condition as between subjects and Algorithm as within subjects factors.

Materials. Three algorithms were used in the experiment. The first one, Lag2, replicated the program described in the previous section. In this case, however, we implemented a Passive Lag2 algorithm which updated the net weights by assigning +1 to wins, 0 to ties and -1 to losses. The second algorithm, Random, played according to the optimal strategy by choosing its moves randomly from a uniform distribution. The third one, Biased, made also random moves but it sampled from a distribution in which one of the moves had a 50% probability of being selected, a second one a 35% probability and the third one 15%. At the beginning of each task the computer assigned randomly one move to each probability class.

Procedure. The experimental sessions were held individually. Participants were instructed about the game rules and it was stressed that each round would be played against a different opponent (in the Classic and Implicit condition) or a different program (in the Nature condition) which followed its own criteria in choosing the moves. After reading the instructions, participants were involved in three 100-turn RPS rounds, each round played against a different algorithm. Participants made their choices by clicking on an image displayed at the vertex of an imaginary equilateral triangle. The images were randomly placed at the vertices for each participant. After participants made their choice, the move played by the computer and the outcome score were shown in the Classic and Implicit conditions while in the Nature condition the scores that could have been obtained by choosing the alternative moves were also displayed. During the task participants were kept informed through a colored bar of their running total that was however reset after each round.

Results

We first ascertained whether the rounds, per se, could influence the participants' performance, i.e., whether the mere fact of having played 100, 200 or 300 RPS turns, independently of the condition and the algorithm, could represent a significant factor in determining their behavior. The scores obtained in the rounds were as follows: *M*=-0.17, *SD*=9.78 for Round1, *M*=1.64, *SD*=11.18 for Round2, and *M*=0.27, *SD*=10.40 for Round3, respectively. A repeated measures one-way ANOVA on the scores of each round did not revealed (*p*=.65) any significant effect. No signs of learning (or fatigue) were thus evidenced that could hinder the interpretation of further results.

We then analyzed the factors manipulated in the experiment, i.e., the Condition to which participants were assigned and the Algorithm against which they played. Table 1 reports the means and the standard deviations of participants' scores.

Table 1: Means (and standard deviations) of the scores.

	Lag2	Random	Biased
Classic	**-3.80** (9.61)	**-1.40** (9.84)	**6.80** (7.14)
Implicit	**-6.45** (10.17)	**-1.20** (5.74)	**5.95** (11.73)
Nature	**-6.40** (9.29)	**4.65** (8.76)	**6.55** (9.60)

A mixed design ANOVA revealed as significant the effect of the Algorithm only ($F(2)=25.13$ $p<.000001$), while Condition ($p=.44$) and the interaction ($p=.29$) did not seem to play any role. In other words, participants behave in the same way when they played the classic RPS game, knowing the relationship that existed between the moves, in the implicit RPS, when the payoff matrix was unknown, and in the non-competitive Nature condition. Two further ANOVAs restricted on the scores of the first 20 and 40 turns, respectively, provided similar results—i.e. the only significant effect was that of the Algorithm—suggesting that no payoff learning was needed to perform in the Implicit condition as in the Classic one, with the results of both conditions were similar to those obtained in the Nature one.

Figure 1 reports the results obtained by collapsing the three conditions. As evidenced, participants lost against Passive Lag2, tied against Random and won against Biased. Two *t* tests confirmed that both in the case of Lag2 (t(t(59)=-4.47, p < .0001) and of the Biased algorithms (t(59) = 5.24, p<.00001) the participants performance differed significantly from 0.

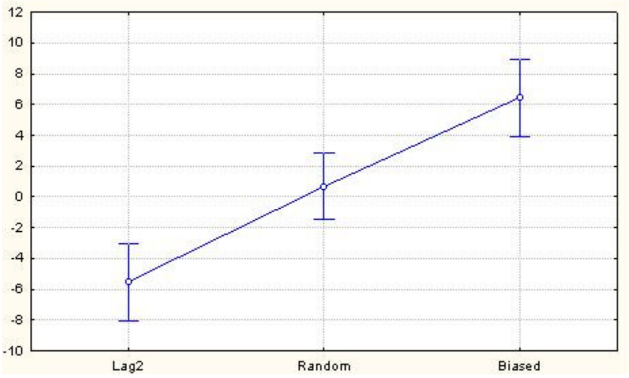

Figure 1. Mean scores obtained against the different algorithms (collapsed conditions) with bars representing 95% confidence intervals.

Modeling the results

Our results contrast with the commonly held assumption that players try to succeed at RPS by anticipating the opponent's move and by making a move capable of beating it. This approach could possibly work in the Classic and Implicit conditions but cannot be applied in the Nature one where the idea of "opponent's move" does not simply make sense. To explain the experimental results we would therefore be obliged to assume that different strategies were applied in the competitive and non-competitive settings. Moreover, following the same assumption, we would expect to find a difference, at least in the first phases, between the Classic and the Implicit conditions. While participants in the former know immediately what to play to defeat an anticipated move, those in the latter have to learn what beats what; since the very first turns, however, the results obtained in these conditions are not discriminable and, again, are similar to those obtained in the third one. Finally, because human participants, as reported by West & Lebiere (2001), had a tough time playing against the Aggressive Lag2 algorithm, we tried to facilitate their task by having them compete against a more manageable version. If Aggressive Lag2, which systematically beats the passive version, represents however the best incarnation of the above mentioned assumption, it is difficult to explain why participants systematically, and independently of any fatigue sign, lost against the Passive Lag2 algorithm.

A principle of economy suggests the possibility that, at least under the conditions examined in our study, participants do not follow different strategies and *do not* try to anticipate the opponent's move but they simply make those moves that are more likely to succeed, independently of the condition to which they have been assigned and the opponent they competed with.

To test this hypothesis, we pitted against our algorithms three different models, representing possible participants' strategies in RPS. The models were the Passive Lag2, Passive Lag1 and a procedural model which exploited the ACT-R's new utility learning mechanism (Anderson, 2007).

The idea on which Procedural ACT-R is based is that an organism, facing the problem of choosing among different moves or actions, will select that which worked best in the past, i.e., the action that was most useful in the previous history of the organism. The model associates therefore to each option a utility measure that is updated after each move application according to the reward it receives from the environment, Starting at an initial value of 0 the utility U_i of each move i at time n is updated according to the formula (Anderson, 2007, p. 160):

$$U_i(n) = U_i(n-1) + \propto [R_i(n) + U_i(n-1)]$$

where α is a learning parameter and R is the outcome received in each turn (the usual +1 for a win, 0 for a tie, and -1 for a loss). The choice of the move is however not deterministic but subjected to noise. The probability P_i that a given move i will be selected among a set of j options (including i too) is given by:

$$P_i = \frac{e^{U_i/s}}{\sum_j e^{U_j/s}}$$

where s it the noise parameter.

The choices made by the model are thus regulated by α and s; we set α to 0.2 and varied s to fit the experimental data. To allow a fair comparison, we implemented NoisyLag2 and NoisyLag1, the nondeterministic versions of the respective models in which the choices were made according to the same formula used in Procedural ACT-R.

A final problem had to be solved before running the simulations. While Procedural ACT-R could be employed in all the experimental conditions, it was not immediately clear how NoisyLag2 and NoisyLag1 could be used to simulate the participant's behavior in the game against Nature, in which the opponent's moves were not available to them. The only data the models had available were represented by the scores obtained by making the different moves. Discarding obviously the idea of storing the score associated with the move chosen by the player, we tried the other options obtaining a surprising result (at least for us): the behavior of the models was exactly the same independent of the fact they stored the best moves (i.e. the moves leading to a score of +1), the worst ones (-1) or those in-between (with a 0 score). In fact, after a moment's thought, we realized that the opponent's moves stored by the West & Lebiere's (2001) lag models were exactly the moves a player should have made to tie in each turn!

That said, we ran a series of simulations with each model starting with s=0.1 and progressively augmenting the parameter through increments of 0.1 up to a final value of 14.0. We simulated 1000 runs of the model for each parameter value against each algorithm, and considered that the model was fitting the data when the 95% confidence intervals of the models' results were completely included within the 95% confidence intervals of the participants' data. The Procedural ACT-R model (with noise values ranging from 0.39 to 0.44) was the only model capable of replicating the participants' performance against all the different algorithms both in term of general performance (total means) and in terms of a temporal series of five successive 20-turns blocks. Both NoisyLag1 and NoisyLag2 did not to fit the participants' data against Lag2 because they were not able, for any s setting, to generate scores that were less than those obtained by the opponent. These models were in a sense too powerful to be considered as a good representation of the people's performance.

Conclusions

In the paper we presented the first results of a research project aimed at investigating the possibility of applying the principle of selection by consequence, traditionally adopted to explain human behavior in games against nature, to model the players' performance in competitive games. We focused on RPS which was previously explained by adopting some form of belief models, i.e. models that "starts with the premise that players keep track of the history of previous play by other players and form some belief about what others will do in the future based on past observation. Then they tend to choose a best-response, a strategy that maximizes their expected payoffs given the beliefs they formed." (Camerer & Ho, 1999, p.2) We found that two models of this kind (NoisyLag2 and NoisyLag1) were isomorphic with models that work by taking into account only the environmental rewards and we found that they were too powerful to be able to explain the human behavior. A purely procedural model based on the ACT-R new utility mechanism was able to fit the experimental data providing thus a simpler and more general explanation for the players' behavior.

References

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought.* Mahwah, NJ: Erlbaum.

Camerer, C., & Ho, T. (1999). Experience-weighted attraction learning in normal form games. *Econometrica, 67,* 827-874.

Fum, D., Napoli, A., & Stocco, A. (2008). Somatic markers and frequency effects: Does emotion really play a role on decision making in the gambling task? *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, 1203-1208.

Fum, D., & Stocco, A. (2003). Outcome evaluation and procedural knowledge in implicit learning. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society.* Boston, MA: Cognitive Science Society, 426-431.

Lebiere, C., & West, R. L. (1999). A dynamic ACT-R model of simple games. *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society,* 296-301. Mahwah, NJ: Erlbaum.

Rapoport, A. & Budescu, D. V. (1997). Randomization in individual choice behavior. *Psychological Review, 104,* 603-617.

Rutledge-Taylor, M. F., & West, R. L. (2004). Cognitive modeling versus game theory: Why cognition matters. *Proceedings of the Sixth International Conference on Cognitive Modeling* 255-260. Pittsburgh, PA: Carnegie Mellong

Rutledge-Taylor, M. F. & West, R. L. (2005). ACT-R versus neural networks in Rock = 2 Paper Rock Scissors. *Proceeedings of the Twelfth Annual ACT-R Workshop, Trieste,* Italy: Università degli Studi di Trieste, 19-23.

Skinner, B. F. (1938). *The behavior of organisms.* New York: Appleton-Century-Crofts.

Stocco, A., Fum, D., & Napoli, A. (2009). Dissociable processes underlying decisions in the Iowa Gambling Task: A new integrative framework. *Behavioral and Brain Functions, 5*, 1

Sutton, R. S., & Bartho, A. G. (1998). *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press.

Thorndike, E. L. (1898). Animal Intelligence; an experimental study of the associative processes in animals. *Psychological Review Monograph Supplement 2*(8):1-109.

Von Neumann, J., & Morgenster, O. (1953). *Theory of games and economic behavior.* Princeton, NJ: Princeton University Press.

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of the literature. *Psychological Bulletin, 77,* 65-72.

Walker, D. & Walker, G. (2004). *The official rock paper scissors strategy guide.* New York: Simon & Shuster.

West, R. L. (1999). Simple games as dynamic distributed systems. *Proceedings of the Complex Games Workshop.* Tsukuba, Japan: Electrotechnical Laboratory Machine Inference Group.

West, R. L. & Lebiere, C. (2001). Simple games as dynamic, coupled systems: Randomness and other emergent properties. *Journal of Cognitive Systems Research, 1,* 221-239.

West, R. L., Stewart, T. C., Lebiere, C., Chandrasekharan, S. (2005). Stochastic Resonance in Human Cognition: ACT-R Versus Game Theory, Associative Neural Networks, Recursive Neural Networks, Q-Learning, and Humans. *Proceedings of the 27th Annual Conference of the Cognitive Science Society,* Mahwah, NJ: Erlbaum, 2353-2358.

Zagare, F. C. (1984). *Game Theory: Concepts and applications.* Newbury Park, CA: Sage Publications, Inc.