Acknowledgements. The work presented here was carried out in the project "Concept Dynamics", which is part of the CRC 360 "Situated Artificial Communicators" at the University of Bielefeld. The research centre is supported by the German Research Foundation. We would like to thank Eric Pfennig for experimental and statistical support, as well as Grainne Delany, Marc Kupietz and Lorenz Sichelschmidt for helpful comments on earlier versions of this paper.

6. References

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*, 115-147.

De Vries, P. H. & Dalenoort, G. J. (1995), The binding problem in distributed systems. In: B. Kappen & S. Gielen (eds.), *Neural networks: Artificial intelligence and industrial applications* (pp. 77-80). Berlin: Springer.

Edelman, S. (1997). Computational theories of object recognition. *Trends in Cognitive Science, 1*, 296-304.

Halford, G. S., Wilson, W. H. & Philips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental and cognitive psychology. *Behavioral and Brain Sciences, 21 (6)*, 803-864.

Henderson, J. M., Pollatsek, A. & Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception and Psychophysics, 45*, 196-208.

Hummel, J. E. & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99*, 480-517.

Kessler, K. & Rickheit, G. (1999). Dynamische Konzeptgenerierung in konnektionistischen Netzen: Begriffsklärung, Modellvorstellungen zur Szenenrekonstruktion und experimentelle Ergebnisse. *Kognitionswissenschaft, 8* (2), 74-96.

Jacobs, R. A. & Kosslyn, S. M. (1994). Encoding shape and spatial relations: the role of receptive field size in coordinating complementary representations. *Cognitive Science, 18*, 361-386.

Marr, D. & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London B. 200*, S. 269-294.

McClelland, J. L. & Rumelhart, D. E. (1988). Interactive activation and competition. In: J. L. McClelland & D. E. Rumelhart (eds.), *Explorations in parallel distributed processing* (pp 11-49). Cambridge, MA: MIT Press.

Miller G. A., (1956). The magical number seven, plus or minus two. Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Schade, U. (1992). *Konnektionismus: Zur Modellierung der Sprachproduktion*. Opladen: Westdeutscher Verlag.

Tarr, M. J. & Bülthoff, H. H. (1995). Is human object recognition better described by geon-structural-descriptions or by multiple-views? Comment on Biederman and Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1494-1505..

Wallis, G. & Bülthoff, H. (1999). Learning to recognize objects. *Trends in Cognitive Science, 3 (1)*, 22-31.

A Memory-based Account of the Prisoner's Dilemma and Other 2x2 Games

Christian Lebiere
Psychology Department
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
cl+@cmu.edu

Dieter Wallach
Institut für Psychologie
Universitaet Basel
Bernoullistr. 16
CH-4056 Basel
dieter.wallach@unibas.ch

Robert L. West
Department of Psychology
Carleton University
1125 Colonel By Drive
Ottawa, Canada K1S 5B6
robert_west@carleton.ca

Abstract

This paper presents a model of the classic game of the Prisoner's Dilemma and related games of strategy. Rather than to rely on the concept of utility central to classical game theory, this cognitive model is based on the theory of human memory implemented in the ACT-R architecture (Anderson, 1993; Anderson & Lebiere, 1998). The model works by adopting the strategy with the best likely expected outcome, as recalled from past experience. The effects of practice and stochasticity in human memory provide an excellent fit to the outcome probabilities and transient behavior found in the empirical data of Rapoport et al. (1976). The model is validated further by applying it unchanged to a set of related games.

Introduction

Games, even the fairly abstract, stylized games studied by game theory, capture many fundamental aspects of human cognition, including problem-solving, meta-cognition, mental imagery as well as more social aspects such as selfishness, altruism and competition. Classical game theory (von Neumann & Morgenstern, 1944; Nash, 1950) defines the solutions to games in terms of fundamental concepts such as equilibrium points, but does not specify how players can learn to reach those points, or even whether they always do. Indeed, Simon (1972) proposed the theory of bounded rationality that asserts that humans do not always behave in such optimal fashion, and Erev and Roth (1998) have proposed that human players are best viewed as cognitively limited learners.

Since economic behavior in general and game playing in particular can be viewed as instances of cognitive behavior, it seems natural to apply general theories of cognition to the study of game theory. Bracht, Lebiere, & Wallach (1998) applied the ACT-R hybrid activation-based production system theory (Anderson, 1993; Anderson & Lebiere, 1998) to the empirical results of the 2-player games modeled by Erev & Roth (1998). Their ACT-R model defined a production associated to each move available to the players and used the expected utility quantities learned for each production to determine which move to make. Lebiere & West (1999) proposed a different ACT-R model of the popular 2-player zero-sum game of Paper Rock Scissors. That model relies on the activation learning mechanisms of the architecture to determine which of the moves, represented by declarative chunks, the opponent is most likely to make next. In this paper, we propose to unify these approaches and apply the resulting model to the classic game of the Prisoner's

Dilemma, which has been seen as an abstraction for economic and social behavior from antitrust litigation to conflict mediation.

The Prisoner's Dilemma

The Prisoner's Dilemma is an instance in the taxonomy of 2x2 games presented by Rapoport et al. (1976) involving 2 players, each having 2 possible moves (or strategies) to choose from. The variant of the Prisoner's Dilemma used by Rapoport is defined by the following payoff matrix:

Player 1 \ Player 2	A	B
A	-1 -1	-10 10
B	10 -10	1 1

Figure 1: Payoff Matrix of the Prisoner's Dilemma.

The first and second rows correspond to the first player's choice of the first and second moves, referred to respectively as A1 and B1. The first and second columns correspond to the second player's choice of the first and second moves referred to as A2 and B2 respectively. The bottom left and top right numbers in each cell define the payoff of that move combination for the first and second players.

The Prisoner's Dilemma is a symmetrical game with respect to the players and is usually played iteratively between the same pair of players. Its fundamental characteristic is that the short-term interest of each player conflicts with their long-term interests. The move combination that maximizes the combined payoffs to the players corresponds to each player selecting the second move also known as the cooperative strategy. For any given iteration of the game, however, the best strategy is always to select the first move, also known as the defecting strategy, because whether the opponent chooses the first or second move it maximizes one's payoff for that iteration. It also has the effect of minimizing the opponent's payoff, which might lead to a retaliation in kind, since the game is symmetrical, leading both players to repeatedly defect and obtain negative payoffs (-1) that are lower than the positive payoffs (1) that could be obtained if they both adopted the cooperative strategy. In game theory parlance, the mutually defecting combination is the natural outcome since it is the only equilibrium, i.e. combination of moves in which neither player can improve its payoff by unilaterally changing its strategy. This equilibrium is called Pareto-deficient because both players are settling for lower payoffs than they could consistently achieve.

Rapoport et al. report the results of running 10 pairs of subjects for 300 plays each. Table 1(a) presents the frequencies of the four possible outcomes for each pair of subjects followed by the frequency average over the 10 pairs. The results are strongly bimodal. 6 of the 10 pairs exhibit what can be described as cooperating behavior, choosing the cooperative outcome (B1B2) for two thirds of the plays or more. 3 of the 10 pairs exhibit

non-cooperating behavior, choosing the defecting outcome (A1A2) for more than half of the time. The percentage of choice of the non-symmetrical outcomes, (A1B2) and (A2B1), is fairly low, averaging 7% and 8% respectively, except in pair #8, in which the (B1A2) outcome is chosen more than half of the time. Overall, subjects choose the cooperative outcome more than half the time (55%) and the mutually defecting outcome less than a third of the time (30%).

Table 1: Frequencies of the four outcomes in the Prisoner's Dilemma.

(a) Data

Subject Pair	A1A2	A1B2	B1A2	B1B2
1	1	1	1	97
2	7	1	1	92
3	14	1	2	83
4	04	5	5	86
5	21	4	3	72
6	24	5	5	66
7	54	12	7	27
8	34	2	52	11
9	58	25	5	12
10	83	9	4	3
Mean	30	7	8	55

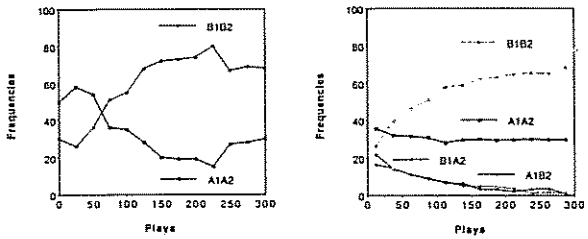
(b) Model

Model Run	A1A2	A1B2	B1A2	B1B2
1	10	13	12	65
2	1	0	2	97
3	4	19	12	65
4	92	4	3	1
5	93	3	3	1
6	1	1	2	96
7	95	3	2	0
8	13	21	18	48
9	2	9	2	87
10	5	4	10	81
Mean	32	8	6	54

One can also analyze the subject's frequencies of choice over time to determine how cooperative behavior arises with experience. Figure 2(a) plots the frequencies of choice of the cooperative (A1A2) and defecting (B1B2) outcome over time. For about the first 50 plays, the defecting outcome seems indeed to be the natural one, being chosen more than half of the time. Its frequency then drops gradually down to about 15% of the time before increasing again and stabilizing at about 30% of the choices by the end of the trial. Conversely, the cooperative outcome is initially chosen about 30% of the time, then after about 50 plays increases sharply to stabilize at about 70% of the choices by the end.

The principal difficulty in modeling the Prisoner's Dilemma is to account for this gradual move away from the equilibrium of the natural defecting outcome and toward the cooperative outcome despite the fact that it is not an equilibrium because of the constant temptation of unilaterally defecting. While the average payoff of the defecting outcome is negative (-1), a player who unilaterally starts to cooperate would see his or her payoff become much more negative (-10) thereby discouraging cooperation, while the other player's payoff for defecting would suddenly increase to the maximum payoff (10) thereby encouraging defection. Any model based on the expected utilities of each strategy would therefore have fundamental problems in accounting for this shift in strategy. It is generally thought that a solution to the Prisoner's Dilemma must involve explicit consideration by the players of the long-term implication of their strategies and the need to cooperate to improve their total payoffs. The ACT-R model presented in the following section demonstrates that this is not necessarily the case and that relying on fundamental characteristics of the

memory system such as stochasticity and adaptivity can provide an explanation for the strategy shift.



(a) Data (b) Model

Figure 2: Frequencies of outcomes over time in Prisoner's Dilemma.

Model

The model presented here is a variation of the model of the simple game of Paper Rock Scissors (PRS) proposed by Lebiere & West (1999), which was itself a variant of the more general sequence learning model of Lebiere & Wallach (1998). Paper Rock Scissors is a 2x3 game in which two players have three possible moves: Paper, Rock and Scissors. The players reveal their move simultaneously, usually by using hand gestures, and the winner is determined by a circular ordering of the moves: Paper beats Rock which beats Scissors which beats Paper. If the opponents choose the same move then a tie is declared. As in the Prisoner's Dilemma, the opponents play each other repeatedly, resulting in a sequence of moves. The PRS model works by trying to guess the opponent's next move given its past moves (usually the last two) and then play the move that beats it. This strategy is implemented in the form of a single production rule:

Paper Rock Scissors
IF the opponent last played moves L2 and L1
  and moves L2 and L1 are most likely followed by move L
  and move L is beaten by move M
THEN make move M, note opponent's move and push a new goal to make the next play

The crucial part of the rule is the second line, which predicts the opponent's next move, L, based on its last two moves L2 and L1. Each encountered move sequence is encoded by a memory chunk and the chunk with the highest activation is retrieved from memory as the basis to guess the opponent's next move. The base-level activation of a chunk (ignoring for simplicity's sake the associative and matching components of the activation formula which are not relevant here) is a combination of adaptivity and stochasticity given by the following formula:

$$A = \ln \sum_{i=1}^{n} t_i^{-d} + N(0, \frac{\pi \cdot s}{\sqrt{3}}) \approx \ln \frac{n \cdot L^{-d}}{1-d} + N(0, \frac{\pi \cdot s}{\sqrt{3}})$$

The first term of the sum represents the strengthening of the chunk activation each time it is (re)created or retrieved from memory. t_i is the time since the ith reference, n is the total number of references and d is the decay rate. Thus, that sum increases with each reference and decreases with time. It can be shown that this provides an account of the power law of practice and power law decay of memory (Newell & Rosenbloom, 1981; Rubin & Wenzel, 1996). If the references are assumed to be evenly distributed over the life of the chunk, then that sum can be approximated as shown in the last part of the equation as a function of n, d and the total life of the chunk L. This is the adaptivity part of the activation equation: chunks that are needed often become more active, and thus more available. Stochasticity is provided by the second term of the sum, normally distributed noise of mean 0 and standard deviation determined by the activation noise parameter s.

For the game of Paper Rock Scissors, this activation equation provides an excellent model both of the subject's constant adaptivity to the opponent's moves as well as the high stochasticity of the results. However, applying this model directly to the Prisoner's Dilemma to try to anticipate the opponent's next move and maximize one's immediate payoff would not provide the needed account, because irrespectively of whether the opponent is going to cooperate or defect, the move yielding the highest immediate payoff is to defect. This reasoning doesn't take into account the long-term cooperation aspect of the Prisoner's Dilemma which is absent in a zero-sum game such as Paper Rock Scissors. However, a straightforward modification of the PRS model yields a more promising account. Instead of trying to determine the opponent's most likely move and trying to defeat it, the model can look at its two possible moves, determine the most likely outcome given each move, and make the move with the best likely outcome. This logic can be implemented by a single production:

Prisoner's Dilemma
IF the goal is to solve the prisoner's dilemma
  and the most likely outcome of making move A is PayoffA
  and the most likely outcome of making move B is PayoffB
THEN make the move associated with the largest of PayoffA and PayoffB,
    note the opponent's move, and push a new goal to make the next play

The most likely outcomes are determined by retrieving the most active of the possible move combinations. Those are represented by chunks of type *outcome* with four slots, *move1* and *move2* which encode the moves of players 1 and 2 respectively and *payoff1* and *payoff2* which encode their respective payoffs. The four chunks necessary to encode the possible outcomes defined by the payoff matrix given in Figure 1 are:

(A1-A2 isa outcome move1 A move2 A payoff1 -1 payoff2 -1)
(A1-B2 isa outcome move1 A move2 B payoff1 10 payoff2 -10)
(B1-A2 isa outcome move1 B move2 A payoff1 -10 payoff2 10)
(B1-B2 isa outcome move1 B move2 B payoff1 1 payoff2 1)

Thus for player1, the first clause of the production will retrieve either of the two chunks associated with player1 making move A, i.e. A1-A2 or A1-B2, and the second clause will retrieve either of the two chunks associated with player1 making move B, i.e. B1-A2 or B1-B2. The payoffs of each move for player1 can be retrieved from slot payoff1, and the higher payoff of the two chunks retrieved (one for move A and one for move B) will determine the move selection. By systematically selecting the largest payoff of the two moves, this model still attempts to maximize its own payoff and does not depend on dubious assumptions of altruism or fairness on the subjects part. But by taking into account the past frequencies of occurrence of each outcome as represented in the activation of the chunks, this very simple model is able to rise above short-term greed and take into account the other player's reaction to one's choices over the course of repeated play. In particular, when the most likely outcome for move A is the mutually defecting outcome A1-A2 with a payoff of −1, if the activation of the chunk B1-B2 encoding the mutually cooperative outcome is higher than the activation of the chunk encoding the outcome when the player cooperates and the other defects, either because of higher past frequency or temporary stochastic fluctuation, then the mutually cooperating outcome with a payoff of 1 will be prefered over the temptation to defect. Since that behavior is self-reinforcing, with the chunk for the mutually cooperating outcome increasing in activation as it is referenced, it would make that outcome a stable choice despite the constant temptation to defect, because the frequencies of the chunks associated with the defecting outcomes encode the fact that the unilaterally defecting outcome with the maximum payoff of 10 is much less likely over the long term than the mutually defecting one with the payoff of -1. Of course, if unilateral defection is initially more common than the mutually cooperative outcome, the latter might not be able to establish itself and will not become a stable outcome. Activation stochasticity will be determine early on which path is taken, but it matters much less as past frequencies become dominant and the game falls into a stable outcome pattern.

The behavior of this model is not fundamentally determined by its parameters. The relevant architectural parameters are the decay rate d, which is left at the standard value of 0.5 used in almost all ACT-R models (Anderson & Lebiere, 1998), and the noise level s, which is set at the value of 0.25, the same as Lebiere & West (1999), Lebiere & Wallach (1998) and Lebiere (1998). A parameter specific to this model is the initial number of references to each outcome chunk. It was set at 10 to reflect a thorough study of the payoff matrix consistent with the experimental conditions of Rapoport et al. Table 1 (b) presents the frequency of outcomes obtained for a sample of 10 runs of the model, to enable a direct comparison with the empirical results. This sample was chosen because as with the subject results it displays 6 cooperative runs, 3 defective runs and a mixed run when no outcome can garner a majority of frequencies. The average frequency of each outcome over the 10 runs is within 2% of the subject frequencies. The specific distribution of individual outcomes cannot of course be guaranteed for each sample of 10 runs because of the bimodal nature of the outcome. Some samples might have 5 cooperations and 5 defections, some 8 cooperations and 2 defections, etc. But the sample run presented here is as representative as Rapoport's group of 10 subject pairs, which he acknowledges is too small to draw definite conclusions on the relative frequency of outcomes. The distribution of outcomes over time is also reproduced by the model, as plotted in Figure 2(b), though not quite as closely. The mutually defecting outcome (A1A2) is initially the most common, but gradually decreases to about 30%. The cooperative outcome is initially less common at

about 30%, but quickly increases to about 60% and then slowly to about 70%. The mixed outcomes, with one player cooperating and the other defecting, gradually decline from about 20% to almost nothing as the model settles into either of the symmetrical outcomes.

Generalization to Other Games

While the very simple model presented above provides a very good fit to the Prisoner's Dilemma data reported by Rapoport et al., a crucial factor in validating a model is determining its generality. Is the model an ad hoc fit to a specific situation or does it capture a fundamental cognitive approach? Fortunately, the Prisoner's Dilemma is only the best known example of a systematic taxonomy of 2x2 games investigated by Rapoport et al. (1976). To determine the generality of the model, we ran it unchanged (other than for the slots values of the outcome chunks encoding the game-specific payoffs, of course) on the ten other games presented in the other chapters of Part 2 of the book. Table 2 presents the payoff matrices of the 10 games and Table 3 presents the frequency of the four outcomes averaged over 10 pairs of subjects. The only changes were the actual payoffs specific to each game.

Table 2: Payoff matrices of the 10 other games.

Game Payoff	A1A2	A1B2	B1A2	B1B2
1	(28 28)	(12 20)	(20 12)	(-10 -10)
2	(28 28)	(-10 20)	(20 -10)	(12 12)
3	(12 12)	(15 21)	(21 15)	(-5 -5)
4	(8 8)	(24 14)	(14 24)	(-5 -5)
5	(7 20)	(16 -5)	(-5 16)	(20 2)
6	(10 19)	(4 -7)	(-7 4)	(19 10)
7	(12 12)	(25 -4)	(-4 18)	(18 25)
8	(15 5)	(5 -10)	(26 16)	(-10 26)
9	(15 20)	(20 15)	(-2 6)	(8 -2)
10	(8 20)	(20 15)	(-2 6)	(15 -2)

Table 3: Frequency of the four outcomes for the other 10 games.

(a) Data

Game	A1A2	A1B2	B1A2	B1B2
1	92	3	3	2
2	63	8	7	22
3	6	50	32	12
4	21	39	39	1
5	80	3	8	9
6	49	4	17	30
7	37	6	5	53
8	30	10	50	10
9	70	14	13	3
10	62	24	9	4

(b) Model

Model	A1A2	A1B2	B1A2	B1B2
1	73	7	7	13
2	43	4	3	50
3	29	28	27	16
4	36	28	31	5
5	77	0	23	0
6	71	5	16	8
7	58	8	7	27
8	30	23	38	9
9	94	6	0	0
10	65	11	23	0

The limitations on the length of this paper prevent us from discussing these games in detail, but the title of the chapters (two related games are presented per chapter) is indicative of the variety of behaviors ascribed to the players in those game situations (Chapter 8: Two No-Conflict Games; Chapter 9: Competition and Cooperation; Chapter 10: Preemption and Procrastination; Chapter 11: A Stable and A Vulnerable Natural Outcome; Chapter 12: Two Forbearance Games; Chapter 13: Two Threat Games). Most but not all of the patterns of the distributions of outcome are reproduced. The correlation between the subjects and the model frequencies is 0.825. Although far from perfect, it is a pretty good indication of the generality of the model given the wide disparity of games and the relatively noisy nature of the results.

Discussion

The ACT-R model of the Prisoner's Dilemma presented here fundamentally relies on the adaptivity and stochasticity of the memory system to determine which outcomes are the most likely to maximize one's long-term payoff. ACT-R's declarative memory tries to adapt itself to the structure of the environment by raising the activation of memory chunks upon use, thus making them more available in the future. Interestingly, this reinforcement occurs upon two separate classes of events which, by default, are treated equally in ACT-R. The first is chunk creation and re-creation. Memory chunks are created when a goal is popped and replaced by another one. If an identical goal is popped again, then instead of creating a separate, identical chunk, the existing chunk is reinforced according to the activation equation given earlier. Thus when a specific outcome chunk is popped to focus on the goal of making the next play, the chunk corresponding to that outcome is reinforced. But when the production fires and retrieves the most active outcome chunk for each move, then those chunks are also reinforced, essentially as if they actually occurred. In a sense, ACT-R doesn't make any distinction between recalling an event and experiencing it again. Another characteristic of the model is that the two players are sharing the same memory chunks. Thus, if a player accesses one chunk it will reinforce it for the other as well. Such synchronization might better be achieved by having each player explicitly take the other's point of view to try and see what they would do, as is common in such games. In any case, more research is needed to determine exactly how important these somewhat accidental characteristics of the model are.

Another feature of the model is that it is sensitive to the relative ordering of the payoffs, but not to their actual values. This is not an unusual assumption since it is generally thought that people are more sensitive to ordering than actual values. Indeed, Rapoport et al.'s exhaustive taxonomy of 2x2 games only depends on relative orderings between payoffs because they are thought to fundamentally determine behavior, whereas actual payoff values only quantitatively modulate frequencies of choice but do not qualitatively affect the outcome. While the model presented here shares with the approach of Bracht, Lebiere and Wallach (1998) the basic precept of explicitly weighting the expected payoff of the available strategies, all the features described in this section are important differences between the two models. More research is needed to explore the implications of these differences in further detail.

Conclusion

We present a model of the classic game of the Prisoner's Dilemma. The model is based on the theory of human memory embodied in the ACT-R cognitive architecture. This model is extremely simple, comprising only a single production, four memory chunks and no parameter tuning. The adaptivity and stochasticity of the ACT-R declarative memory are responsible for the model's fit to the empirical data of the distribution of outcome frequencies and their evolution over time. The model is validated by showing that it can provide a good overall fit to the data from multiple different games, without changing the model in any way.

Acknowledgments

This research was supported by a grant from the Office of Naval Research (N00014-95-10223) for the first author.

References

Anderson, J. R. (1993). *Rules of the Mind.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J. R., & Lebiere, C. (1998). *The Atomic Components of Thought.* Mahwah, NJ: Lawrence Erlbaum Associates.

Bracht, J., Lebiere, C., & Wallach, D. (1998). On the need of cognitive game theory: ACT-R in experimental games with unique mixed strategy equilibria. Joint Meetings of the Public Choice Society and the Economic Science Association, New Orleans, LA.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88(4), 848-881.

Lebiere, C. (1998). The dynamics of cognition: An ACT-R model of cognitive arithmetic. *CMU Computer Science Dept Technical Report* CMU-CS-98-186. Pittsburgh, PA.

Lebiere, C., & Wallach, D. (1998). Implicit does not imply procedural: A declarative theory of sequence learning. Paper presented at the *Forty First Conference of the German Psychological Society*, Dresden, Germany.

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

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

Newell, A. & Rosenbloom, P.S. (1981). Mechanisms of skill acquisition and the power law of practice. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp.1-56). Hillsdale, NJ: Lawrence Erlbaum Associates.

Simon, H. A. (1972). Theories of bounded rationality. In C. B. Radner & R. Radner (Eds.), *Decision and Organization* (pp. 161-176) Amsterdam: North-Holland.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior.* Princeton University Press.

Rapoport, A., Guyer, M.J., & Gordon, D. G. (1976). The 2x2 Game. Ann Arbor, MI: University of Michigan Press.

Rubin, D.C. & Wenzel, A.E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological review, 103*, 734-760.