Applying Cognitive Architectures to Decision-Making: How Cognitive Theory and the Equivalence Measure Triumphed in the Technion Prediction Tournament

Terrence C. Stewart (tcstewar@uwaterloo.ca)
Centre for Theoretical Neuroscience, University of Waterloo
Waterloo, ON, N2L 3G1

Robert West (robert_west@carleton.ca)
Institute of Cognitive Science, Carleton University
Ottawa, ON, K1S 5B6

Christian Lebiere (cl@cmu.edu)
Psychology Department, Carnegie Mellon University
Pittsburgh, PA, 15213

Abstract

For the Technion Prediction Tournament, we developed a model of making repeated binary choices between a safe option and a risky option. The model is based on the ACT-R declarative memory system, with the use of the Blending mechanism and sequential dependencies. By using established cognitive theory, rather than specialized machine learning techniques, our model was the most predictive when generalizing to new conditions. However, we did not tweak parameters to minimize prediction error; instead we maximized the number of different conditions producing statistically equivalent behavior. If we had not done this the model would not have won the tournament. This leads to the paradoxical result that by emphasizing cognitive explanation over prediction, we achieve more accurate predictions.

Keywords: decision-making; cognitive modeling; equivalence; ACT-R; blending; sequential dependencies;

Repeated Binary Choice Decisions

The effects of rewards on decision-making are highly studied, and a wide variety of effects have been observed. In the simplest paradigm, two choices are presented to a participant, and once a decision has been made an explicit numerical reward is provided. If this process is repeated many times, participants will start to favor one choice over the other if it is rewarded more.

The standard empirical result is "probability matching". Here, if option **A** has a probability p of providing more reward than option **B**, then participants would choose option **A** with probability p. Interestingly, this is very different from the optimal strategy of choosing **A** if p>0.5 and otherwise choosing **B**. However, Friedman and Massaro (1998) note that "probability matching in binary choice ... is less robust than most psychologists seem to believe."

Many more complex effects have since been identified. For example, in the *Loss Rate Effect*, "when the action that maximizes expected value increases the probability of losses, people tend to avoid it" (Erev & Barron, 2005, p. 917). That is, a choice that has a higher expected value in the long run will be chosen less often if it is comprised of many small losses and few large gains. In the *Payoff Variability Effect*, individuals will switch from risk-seeking to risk-aversion depending on the variance of the reward. However, many of these effects are subject to variations as a function of learning, individual differences, and other factors (Lebiere, Gonzalez & Martin, 2007).

Technion Prediction Tournament

To encourage the creation and evaluation of models of this fundamental component of human decision-making, Ido Erev organized a competitive modeling tournament called the Technion Prediction Tournament. The tournament had three divisions. The division of interest for us involved modeling human behavior in different versions of a repeated binary-choice game. Empirical data was gathered on 120 randomly chosen empirical conditions with different rewards. In each condition, one option always produced the same deterministic reward M, while the other option would produce the reward H with probability p_H and otherwise produce the reward L. Rewards and probabilities were chosen to make the expected value of each choice roughly even, emphasizing attitudes toward risk rather than abilities to estimate reward. For each condition, 20 participants made 100 decisions, receiving a numerical reward after each choice. This type of task is meant to capture the essential qualities of what most people would call a game or competition (e.g., tennis, baseball, boxing, paper-rock-scissors, poker).

The competition also included two other divisions where only a single choice was made after subjects either learned or were told the reward structure. Intuitively, these conditions model informed human decision-making. Neither of these divisions is considered here. For complete details on the tournament, see Erev et al., (in press).

As part of the competition, empirical data on 60 of the 120 conditions was publicly released. Researchers were free to use this data to produce predictive models that were then tested by examining their predictions on the remaining 60 conditions.

The model presented here won the tournament in the repeated game division. That is, it produced more accurate predictions (in terms of mean squared error) on the testing data set than any of the other models in the division. Due to limited space Erev et al., (in press) provides only a brief discussion of the model. However, the model had two unique features that merit further attention. The first is that, unlike the other models in this competition, our model was not a specialized model of game playing. Instead it was a cognitive model of decision-making based on the human memory system. The second is that, unlike the other models in the competition, we used Equivalence Testing (Stewart, 2007; Stewart & West, 2007) to set the model parameters. Equivalence testing emphasizes the degree of *statistical equivalence* between a set of observed empirical measures and the corresponding model outputs (i.e., as opposed to using a regression to get the best possible fit to a data set).

The Decision-Making Model

The basic idea behind our model of making decisions in a repeated binary choice context was to treat it as a memory task. This allowed us to leverage extensive previous research in terms of the performance and accuracy of human memory.

ACT-R Declarative Memory

The core of the model is the declarative memory system from the ACT-R cognitive architecture (Anderson & Lebiere, 1998), which has been used as the basis for a broad range of explicit and implicit recall tasks. The general principle is that the odds of a memory being needed decay as a power law over time, and that, if an item appears more than once, these odds are summed, again as a power law, over all occurrences. This principle is a close match for realistic human cognitive environments (Anderson & Schooler, 1991).

To implement this, each item i in memory is given an activation level A_i, calculated using Equation 1, where t_k is the amount of time since the k^{th} appearance of this item, d is the decay rate, and $\varepsilon(s)$ is a random value chosen from a logistic distribution.

$$A_i = \ln \sum_{k=1}^{n} t_k^{-d} + \varepsilon(s) \qquad (1)$$

When a memory is to be retrieved, the activation level is calculated, and if it is above a retrieval threshold τ then it is successfully retrieved. The amount of time required for recall to occur is given by Equation 2, where F is a latency scaling factor.

$$T = F e^{-A_i} \qquad (2)$$

The ACT-R model of declarative memory has been applied to model two aspects of human game-playing abilities. The first is our somewhat limited ability to learn and exploit probabilities and payoffs. To do this it is necessary to compare *expected* rewards for each choice. For example, if option **A** gives a reward of 8 and option **B** gives a reward of 10 half the time and otherwise a reward of 5, the expected reward for **B** (i.e., the average reward over time) is 7.5. Therefore, **A** should be preferred to **B** (although, as noted above, people will still choose **B** some of the time). The ACT-R declarative memory system, by itself, does not produce this effect. However, Lebiere (1999) created an augmentation called the Blending Mechanism that allows ACT-R to do this. The Blending Mechanism, which is described below, in combination with the ACT-R declarative memory system, has been used successfully to model this type of game (e.g. Sanner et al., 2000; Lebiere et al., 2003; Lebiere, Gonzalez & Martin, 2007).

The second aspect of human game-playing that has been modeled using ACT-R is the human ability to capitalize on sequential dependencies in their opponents' outputs (e.g., West & Lebiere, 2001; West, Lebiere & Bothell, 2006). Although the ACT-R declarative memory system was not designed with this in mind, this ability falls out naturally from the way it works. It requires only that information about previous trials is stored in chunks along with the current outcome. This approach has been successfully used to model the human ability to detect and exploit sequential information, but has not previously been integrated with Blending.

In terms of the competition the Blending model seems most relevant since the task is to learn the probabilities and payoffs, and there are no sequential dependencies to detect. However, we found that Blending combined with detecting sequential dependencies worked better than Blending alone. That is, the search for sequential dependencies where there were none made the model outputs more human-like, suggesting that humans do not turn off this ability when it is not needed. Essentially, the effect of this is to dampen the impact of the memory-based mechanism, especially for recent results and rare results.

Sequential Dependencies in the Model

In ACT-R each memory consists of a set of slot-value pairs and is referred to as a chunk. For this particular model, each chunk consists of which button was pressed, the numerical reward that was received, and the history of button-pressing leading up to the current press. The number of previous button presses was set to 2, as evidence indicates that this setting closely matches human performance (e.g. West et al., 2005). This memory representation is a direct encoding of the relevant information available to the decision-maker in the current context and does not require deliberate cognitive strategies. As such, it suggests a model of implicit decision-making that reflects the constraints of the human architecture rather than design decisions made by the modeler.

An example memory chunk is given in Table 1. This configuration would occur if the participant pressed the right button twice, and then pressed the left button and got a reward of 8.4.

Table 1: Sample memory chunk

Slot	Value
choice	left
reward	8.4
lag_1	right
lag_2	right

When performing a recall, the model only considers memories whose recent history match the current recent history. That is, the chunk shown in Table 1 will only be a candidate for recall when the model has just finished selecting the right-hand button twice in a row.

Blending in the model

When the model attempts to recall a chunk that matches the recent history, multiple chunks may be found. In the competition, when attempting to recall an expectation for the button associated with the risky choice, there will be two chunks in memory: one from previous situations where the Low reward was received and one where the High reward was received. The chunk with the higher activation will be the one that has occurred the most in the past and/or most recently (since the learned chunk activation reflects both recency and frequency effects), and is therefore more likely to contain the correct outcome for the current trial. However, because of the probabilistic nature of the payoff, it may not be the best choice. For instance, it could occasionally lead to very negative consequences that would offset the more common but limited gains. The blending mechanism was developed for this type of situation and has been used on other instance-based learning and decision-making tasks (e.g. Gonzalez et al., 2003).

To blend the two chunks that match the current situation, the numerical value for the rewards are combined using the activation value (Equation 1) as a weighting factor. This results in a blended reward value *r*, as shown in Equation 3. This is then taken as the expected reward.

$$r = \frac{\sum_i A_i r_i}{\sum_i A_i} \quad (3)$$

Algorithm

Given this memory system, the underlying algorithm is straightforward. Two recalls are attempted to get an expected reward for each of the two options. These occur sequentially and in random order. If either recall fails (i.e. if there are no matching previous memories or if their activation is below the retrieval threshold τ), then that option is chosen. This was done to model exploratory behavior, since the participant cannot remember or has never seen the results of choosing that button in that context. If both retrievals succeed, the one with the largest expected reward (Equation 3) is chosen (choosing randomly if they are exactly equal). Once the reward is provided, a corresponding chunk is added into the declarative memory system to reflect what actually happened. Then the history is updated and the system is ready to make the next prediction.

For consistency with the use of the ACT-R declarative memory system, this algorithm was implemented using the ACT-R production system. This expresses each of the above steps using if-then rules, each of which requires 50 milliseconds to occur. Note that this, combined with Equation 2 for determining how long it takes to retrieve a memory, allows the model to give predictions for the time taken to make its decision. This timing information also impacts the performance of the model, since Equation 1 indicates how memories decay over time.

For the competition, the model needed to predict average performance over 100 trials given a particular experimental situation. To create this prediction, 1000 separate models were generated and each one was run through 100 trials. The final prediction was the average proportion of times the risky choice was made.

The source code for this model is available at <http://ccmlab.ca>.

Parameter Exploration

As with any computational model, there are a variety of numerical parameters that can be adjusted. However, since the model is based on the ACT-R cognitive architecture, we can turn to previous experiments to help constrain these model parameters. For example, the parameter *d* in Equation 1 is consistently set to 0.5 in ACT-R models to produce results that are predictive of human performance on many memory tasks. The model parameters are shown in Table 2, along with the standard values used for each one.

Table 2: Canonical parameter values for the model

Parameter	Canonical Value	
decay	0.5	(*d* in Equation 1)
noise	0.3	(*s* in Equation 1)
latency	0.05	(*F* in Equation 2)
retrieval threshold	0	(τ)
lag	2	(size of context)
production time	0.05	(time to apply a rule)

Of these parameters, only two are commonly changed in ACT-R models: noise (*s*), because it can be used to represent both specific retrieval stochasticity a number of other sources of unpredictability, and retrieval threshold (τ), because it is used to compensate for constant variations in other activation factors. However, both parameters were searched over intervals close to their canonical values (0.25 for the noise; 0 for the retrieval threshold). It is thus useful to explore the behavioral changes in the model as these parameters are adjusted.

Equivalence Testing

The Technion Prediction Tournament provided raw empirical data for 20 subjects performing 100 decisions in each of 60 different experimental conditions. Each of these conditions provides a separate measure for evaluating the model's performance.[1]

The standard metric for model quality over a set of measures is the root-mean-squared error (RMSE), often used to find the "best fit" parameter setting for a model. This measure is shown in Figure 1, indicating that the smallest prediction error averaged over the 60 different conditions occurs with a very low noise value (s<0.01) and a retrieval threshold τ of -1.

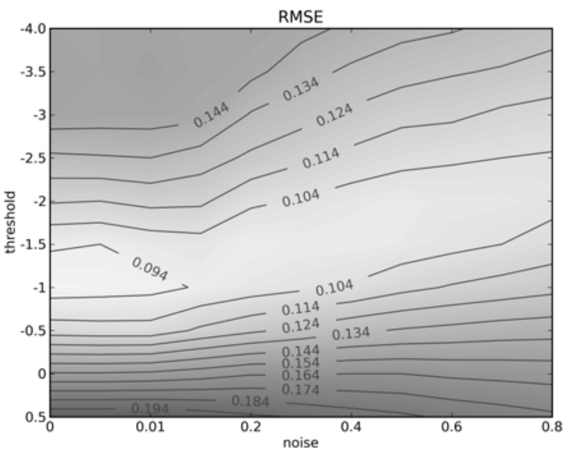

Figure 1: Root Mean Squared Error for the model on the 60 experimental conditions. Each point indicates the RMSE for a different setting of the noise and threshold parameters.

However, the RMSE approach can be difficult to interpret. Firstly, it averages over experimental conditions, meaning that if there are a few conditions for which the model is highly inaccurate, it can still have a small RMSE. Secondly, and more fundamentally, this approach does not take into account sampling error in the empirical data. Given that only 20 participants were used for each condition, the confidence intervals for each measure may be fairly large, making the RMSE approach prone to over-fitting.

To determine the overall quality of the model over all 60 conditions, we did not use the standard approach of minimizing the root-mean-squared error. Instead, relativized equivalence (Stewart & West, 2007) was used. Here, the key measure is the *worst-case equivalence*. Equivalence is defined as the maximum difference between the 95% confidence intervals of the human participants and the model. That is, it is the number for which there is 95% confidence that the human performance and the model performance differ by less than this amount. This approach is derived from the equivalence test (Barker et. al, 2002) used in epidemiology. If a normal distribution is assumed, this range can be reduced (Tryon, 2001), but the work presented does not do this, and instead uses bootstrap confidence intervals (Davidson & Hinkley, 1997) so as to make no assumptions about the distribution of the data.

To determine the relativized equivalence E_r between a particular parameter setting and the participants' performance over the 60 measures provided, Equation 4 was used, where the model's confidence interval on situation i is $M_{i,L}$ to $M_{i,U}$ and the human participants' confidence interval is $H_{i,L}$ to $H_{i,U}$. This gives a result that is normalized so that a value of 1.0 indicates that all model values are within the corresponding confidence intervals (i.e. any model with E_r<1 is not statistically distinguishable from the real participant performance on any particular situation).

$$E_r = max_i \frac{max(M_{i,U} - H_{i,L}, M_{i,U} - H_{i,L})}{H_{i,U} - H_{i,L}} \quad (4)$$

This method was developed to more conservatively characterize the behavior of a model. By focusing on the worst-case scenario (rather than averaging over situations as in MSE approaches), it clarifies that the model is suitable for all of the situations being investigated. The results of this metric are shown in Figure 2.

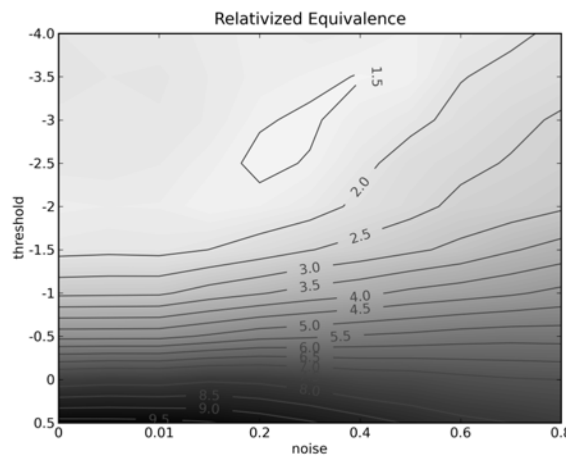

Figure 2: Relativized Equivalence for varying parameter settings.

When making use of this method, it is common to find there are particular conditions in the fitting set in which the model produces a high E_r value regardless of the parameter settings. This is especially true with a large number of conditions: with 60 conditions, even a perfect model will be expected to be outside a 95% confidence interval on three conditions, just by chance. Imperfect models and/or imperfect data further increases the likelihood of this happening.

For this competition, 9 out of 60 conditions in the fitting set were identified as problematic. That is, no parameter settings were found that would give low E_r values on these

[1] In the current work, individual differences were not modeled.

conditions and also maintain low E_r values on the majority of the other conditions. Because this could be due to an unknown bias or outliers in these experimental conditions, they were excluded from our analysis. These conditions are shown in Table 3. Re-running the analysis using a replication, or a conceptual replication, of the same experimental conditions could help identify problematic conditions for the model, if they exist. However, at this time the necessary replications have not been performed.

Table 3: Experimental conditions identified as outliers

#	*H*	p_H	*L*	*M*
7	-5.6	0.7	-20.2	-11.7
13	-2	0.05	-10.4	-9.4
20	-4.3	0.6	-16.1	-4.5
21	2	0.1	-5.7	-4.6
24	9.2	0.05	-9.5	-7.5
30	3	0.91	-7.7	1.4
36	5	0.08	-9.1	-7.9
45	2.8	0.8	1	2.2
49	13.4	0.5	3.8	9.9

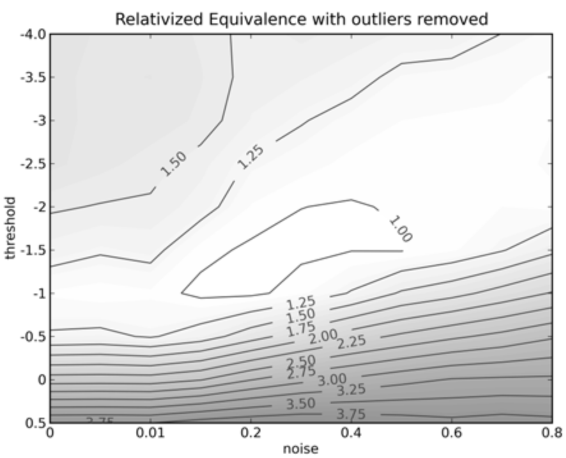

Figure 3: Relativized Equivalence with outlier measures removed.

Once these outliers are remove (as shown in Figure 3), we can see that setting the noise parameter to values between 0.2 and 0.4 and the threshold to values between -1 and -2 produces models that are equivalent to the empirical data. All of these models have E_r values below 1.0. That is, there is no statistically significant difference between the model's behavior and the observed behavior over any the 51 non-outlier measures considered. This indicates that the model is successfully capturing the human behavior at the level that is statistically warranted.

For the purposes of the Technion Prediction Tournament, a single parameter value had to be chosen. We selected the center of the equivalent region, giving a threshold of -1.6 and a noise of 0.35, which is close to the canonical value for noise 0f 0.25 used in previous models of this type.

Generalization

To evaluate the various models in the tournament, their ability to generalize to a testing set was measured. This was done using the standard RMSE approach. Our model won the tournament with a RMSE of 0.087, and the next closest model scored 0.092.

Interestingly, if we had have used the standard best fit approach rather than the Equivalence methodology, we would not have won the tournament. As can be seen in Figure 1, the best fitting model on the training data had a noise of 0.001 and a threshold of -1. However, Figure 4 shows that this model performs considerably worse on the testing data, giving a RMSE of 0.096.

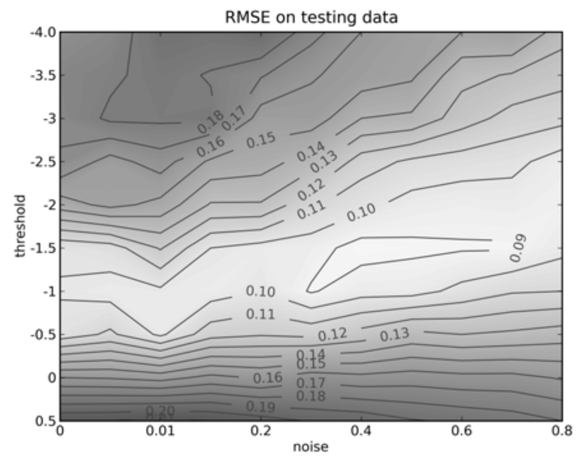

Figure 4: Root Mean Squared Error on the testing data for varying noise and threshold.

Discussion

By building a decision-making model using well-established cognitive models, we successfully predicted behavior in a novel domain. Our model beat a wide variety of machine-learning techniques; the next best models involved two-stage sampling and normalized reinforcement learning. None of the competing models other than ours made use of general knowledge about human cognition. Instead they relied on mathematical optimization techniques. As Lebiere et al. (2003) demonstrated, modeling methods that rely on general assumptions about cognitive invariants, such as cognitive architectures, and can generalize models across a range of paradigms and conditions can be superior to machine learning techniques such as Bayesian networks or Markov models on a number of counts: (a) they require less data to be parameterized because unlike machine learning methods that attack each new problem tabula rasa, constraints inherited from other models prune the parameter space, (b) they require fewer domain-specific assumptions because cognitive constraints constrain the relevant problem representation rather than leaving it entirely to the modeler, and (c) they allow a more complex representation of the problem-solving state, such as a combination of symbolic structures and statistical parameters such as activation.

Conclusions

We have established a novel model of human decision-making in repeated binary choice conditions where one option gives a fixed reward and the other option gives a reward that is randomly selected from two possible values. While this model produces the smallest prediction errors among those entered into the Technion Prediction Tournament, we can also draw a stronger conclusion. In particular, this model produces behavior that is statistically indistinguishable from the human performance, given the available empirical data.

While a few of the 120 experimental conditions did have to be removed in this analysis, this removal does not invalidate the model. Given the large number of conditions, it is expected that even a perfect model would fail to match due to sampling error. To establish whether these conditions do actually indicate problems with the model, more empirical measures are needed. If these measures are consistent with the model, then this is a case of sampling error. If these measures continue to be inconsistent then we will have sufficient evidence to adjust our model to take this into account. However, without further empirical evidence there is no statistical justification for attempting to fit our model more closely to the human performance. By following the equivalence method for evaluation we successfully avoided this over-fitting.

References

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

Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2*, 396-408.

Barker L.E., Luman E.T., McCauley M.M., & Chu Y.R. (2002) Assessing equivalence: An alternative to time use of difference tests for measuring disparities in vaccination coverage. *American J. of Epidemiology, 156*, 1056-1061.

Davison, A.C. and Hinkley, D.V. (1997). Bootstrap Methods and Their Application. Cambridge University.

Erev, I. and Barron, G. (2005). On Adaptation, Maximization, and Reinforcement Learning Among Cognitive Strategies. *Psych. Review, 112*(4), 913-931.

Erev, I., Ert, E., Roth, A., Haruvy, E., Herzog, S., Hau, R., Hertwig, R., Stewart, T., West, R., and Lebiere, C. (in press). A choice prediction competition for choices from experience and from description. *Journal of Behavioral Decision Making: Special edition on Decisions from Experience.*

Friedman D., and Massaro D.W. (1998). Understanding variability in binary and continuous choice. *Psychonomic Bulletin & Review, 5*, 370–389.

Gonzalez, C., Lerch, F. J., & Lebiere, C. (2003). Instance-based learning in real-time dynamic decision making. *Cognitive Science 27* (4), 591-635.

Lebiere, C. (1999). Blending. In *Proceedings of the Sixth ACT-R Workshop.* George Mason University, Fairfax, VA.

Lebiere, C., Gonzalez, C., & Martin, M. (2007). Instance-based decision-making model of repeated binary choice. In *Proceedings of the 8th International Conference on Cognitive Modeling*. Ann Arbor, MI.

Lebiere, C., Gray, R., Salvucci, D., West, R. (2003). Choice and learning under uncertainty: A case study in baseball batting. In *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, 704-709. Mahwah, NJ: Lawrence Erlbaum Associates.

Lebiere, C., & Wallach, D. (2001). Sequence learning in the ACT-R cognitive architecture: Empirical analysis of a hybrid model. In Sun, R. & Giles, L. (Eds.) *Sequence Learning: Paradigms, Algorithms, and Applications*. Springer LNCS/LNAI, Germany.

Lebiere, C., & West, R. L. (1999). A dynamic ACT-R model of simple games. In *Proceedings of the 21st Annual Conference of the Cognitive Science Society*, 296-301. Mahwah, NJ: Erlbaum.

Pylyshyn, Z. W. (1984). *Computation and Cognition: Towards a Foundation for Cognitive Science*. Cambridge: MIT Press.

Sanner, S., Anderson, J. R., Lebiere, C., & Lovett, M. C. (2000). Achieving efficient and cognitively plausible learning in Backgammon. *Proceedings of The Seventeenth International Conference on Machine Learning*. San Francisco: Morgan Kaufmann.

Stewart, T. C. (2007). *A Methodology for Computational Cognitive Modelling.* PhD Thesis, Institute of Cognitive Science, Carleton University.

Stewart, T. C., West, R.L. (2007) Equivalence: A Novel Basis for Model Comparison. In *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*.

Tryon, W. (2001). Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: An integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods, 6*(4), 371-386.

West, R. L., & Lebiere, C. (2001). Simple games as dynamic, coupled systems: Randomness and other emergent properties. *Journal of Cognitive Systems Research, 1(4)*, 221-239.

West, R. L., Lebiere, C. & Bothell, D. J. (2006). Cognitive architectures, game playing and human evolution. In Sun, R. (Ed) *Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation*. NY, NY: Cambridge University Press.

West, R. L., Stewart, T. C., Lebiere, C., & Chandrasekharan, S. (2005). Stochastic resonance in human cognition: ACT-R vs. game theory, associative neural networks, recursive neural networks, Q-learning, and humans. In B. Bara, L. Barsalou & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Lawrence Erlbaum Associates.