Biologically Inspired Cognitive Architectures II:
Papers from the AAAI Fall Symposium (FS-09-01)

Applied Cognitive Models of Frequency-based Decision Making

James J. Staszewski and Christian Lebiere

Department of Psychology, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213
jjs@cmu.edu; cl@cmu.edu

Abstract

In this paper, we present a cognitive model of frequency-based decision-making applied to the task of landmine detection. The model is implemented in the ACT-R cognitive architecture and is strongly constrained by the cognitive primitives of the architecture. We then generalize the model to another task in the domain of macroeconomic decision-making using the same architecture, pursuing theoretical parsimony. We describe each model's representation requirements, assess their fits to the data, and analyze their performance scaling as a function of task and architectural parameters. Efforts to generalize the landmine detection model to macroeconomic decision making showed that reasonable fits to the macro-economic performance data could be achieved by models based either on procedural knowledge or declarative knowledge. This finding underscores the importance of distinguishing between processing strategies employed to execute tasks. Such detail appears needed to understand the neural foundations of frequency-based decision-making.

Introduction

Decision based on the frequency of events is a common occurrence in everyday life. For instance, we might decide whether to take an umbrella to work based on our relative experience of rainy and sunny days. We don't seem to spend much time thinking about it, and certainly don't go to any great length to keep statistics and explicitly estimate probabilities, but that information forms the basis for many of our decisions and the nature of our cognition is reflected in the quality of those decisions.

Apart from scientific interest in frequency information processing, such phenomena have significant practical implications for some frequency-based decisions involve far more than getting wet. For example, frequency-based decision-making has been linked to economic policy making (Blinder and Morgan, 2000) and perception of medical risks (Lipkus, 2007). We will shortly discuss a task in which the quality of frequency-based determines the fate of limbs and lives. Sedlmeier and Betsch (2002) survey a number of conditions that have been found to influence frequency estimation, including spacing, repetition, order, feedback and base rate effects. They also describe competing approaches to frequency processing: those relying on basic memory capacities to automatically store events and emit frequency judgments vs. those emphasizing strategies and heuristics.

We approach the issue of frequency-based decision making from the perspective of a cognitive architecture, ACT-R (Anderson, 2007; Anderson & Lebiere, 1998). Consistent with the goals of theoretical integration, we pursue general theoretical principles that can explain frequency-based decisions by analyzing relevant phenomena in two quite different task domains: landmine detection and a laboratory analog of sequential economic decision-making. We will describe these tasks and their common elements-- decision-making based on the relative frequency of events--and then present two competing models. We will proceed to evaluate those models, their shortcomings and compatibilities, and draw some general lessons for the cognitive modeling enterprise.

Landmine Detection

As argued in the introduction, frequency-based judgments in decision-making occur in a wide range of domains. If one wants to claim having developed a model of the basic skill, it is important that it be applied to different domains to ensure that its success is not due to having successfully reverse-engineered the constraints of a specific domain.

The first domain for which we developed a model of frequency-based decision making is landmine detection. Safe, effective, and efficient mine detection with a state-of-the-art device such as the AN/PSS-14 requires well-trained operators (Staszewski, 2004). Past work designing and developing training for new AN/PSS-14 operators has demonstrated the value of applying an understanding of the thought processes that underlie expert skill for effective training design. However effective the current AN/PSS-14

Program of Instruction (POI) is currently, a recent evaluation of the training program given to US soldiers, has identified areas within the POI where fundamental knowledge is needed to enhance training effectiveness. In particular, the lack of sound, clear, and explicit description of the criteria that operators should use to interpret the output of the ground-penetrating radar (GPR) to ensure detection of buried landmines and enable maximally accurate discrimination of landmines from other non-threatening buried "clutter" objects that nonetheless generate system responses on both the PSS-14's metal detector (MD) and GPR sensors.

Basic detection procedures involve first performing broad sweeps using either sensor to determine the overall location of a buried object. The operator then searches this area with the MD sensor to create a "magnetic halo" and locate the likely position of the object. The operator then executes a series of short back-and-forth sweeps using the GPR sensor. Unlike the MD sensor, which emits a continuous, modulated sound, the GPR sensor outputs a discrete tone when it detects a discontinuity beneath the soil surface. The sequence of tones generated by this series of sweeps is used to classify the object as either mine or clutter making the task a frequency-based decision-making problem.

Declarative Model

In addition to the basic procedure described above, a number of additional aspects are known about the interpretation of GPR signals. One is that it is an iterative process, with the operator refining its judgment with each sweep. As for the previous task, it is also a self-terminating process, with the operator deciding when enough information has been acquired to provide the basis for a clear decision. Also, time pressures (e.g. clearing a minefield in combat situations) often force a tradeoff between gathering more information and making a quick decision.

Gonzalez & Lebiere (2005) described a number of applications of a declarative modeling paradigm: instance-based models of decision-making. Those models share a feature: given a current decision context, they rely on past instances of decisions and their outcomes stored in declarative memory to find the best decision for the current context. Applied to this problem, the model works as follows: at any given point, the previous sequence of signals is summarized by an evaluation state that reflects the current judgment and provides the context for the next decision. During the next sweep, the model will combine that context with the new signal outcome to update the evaluation state. This key step is mediated by retrieval from declarative memory of previous, similar instances encountered during training and their actual outcomes. When the evaluation exceeds a threshold in either direction (i.e. very high, indicating a mine, or very low, indicating a clutter object), the model makes a decision and learns its outcome. Each decision-making instance at every sweep consisting of the context, the signal, and the revised evaluation (now the ground truth) enters declarative memory. The key attribute of declarative memory is its ability to generate an expected evaluation V based on the aggregate of past instances that weighs the evaluation Vi stored by each instance by the saliency of that instance (as reflected in its probability of retrieval Pi) according to the following blending equation:

$$V = \arg\min \sum_i P_i \cdot \left(1 - Sim(V, V_i)\right)^2$$

This process has been used in a range of domains from position evaluation in a masters-level backgammon player (Sanner et al, 2000) to the winning entry in the Technion Prediction Tournament (Stewart, West & Lebiere, 2009).

Functional Analysis

We are still in the process of collecting field and laboratory data and could not compare the model to either expert or trainee performance. However, it is essential to also understand the functional characteristics of the model as it attempts to perform the task. Therefore, we simulated task conditions to the best extent possible. Mines and clutter objects were associated with a 75% and 25% probability of emitting a signal on any given GPR sweep, respectively. We exposed the model to a training period where it practiced its judgment and accumulated decision-making instances for both mine and clutter objects, then tested it without further learning to determine its performance. The number of sweeps for each training instance was capped at 10, but testing runs were allowed to request up to 20 sweeps.

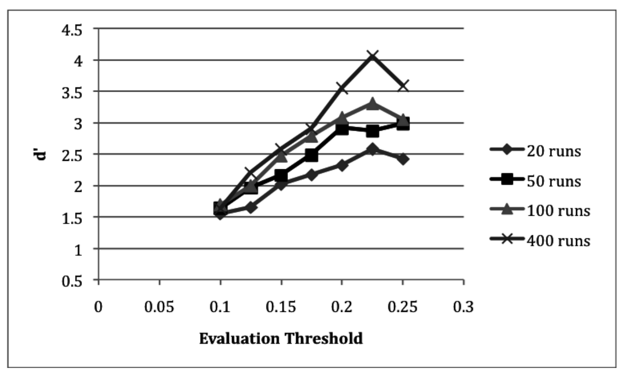

Figure 1: d' as a function of evaluation threshold for various training lengths.

Figure 1 displays the model's detection performance in terms of d' measure as a function of the selectivity of the decision threshold, i.e. the deviation in either direction from the initial evaluation on the evaluation scale that would lead to a decision. One can see that as expected detection performance increases with selectivity, up to a plateau where the threshold is often not reached and a decision is forced upon the model at the end of the 20 sweeps limit. Detection performance also increases as a function of the number of training runs (20, 50, 100, 400) before testing, indicating extended learning.

These basic runs only involve the presence or absence of a signal ("hit") on each sweep. To test the ability of the model to detection statistical patterns beyond straightforward frequency hit frequency, we introduced a number of complexities typical of real-world GPR signals. One is the presence of double signals ("double hits") over some targets. Figure 2 shows that compared to the base condition without double hits ("zero") and a condition in which the probability of a double hit is not target-specific ("even"), the model is able to improve its detection performance when double-hits are disproportionately associated with either mines ("same") or clutter ("fold"). It should be noted that the only change to the model was add another type of signal without giving it any indication of its semantics: it was able to learn it on its own, implicitly.

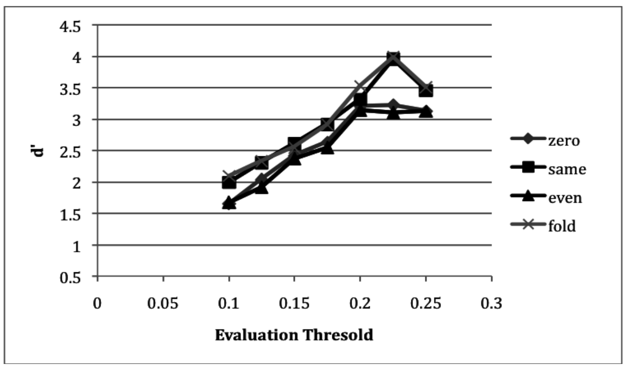

Figure 2: d' as a function of evaluation threshold for various double-hit distributions.

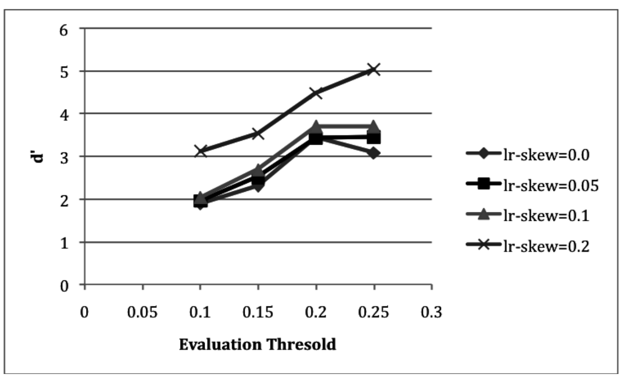

Figure 3: d' as a function of evaluation threshold for various left-right skew factors.

Another manipulation observed in the PSS-14 detector is a differential sensitivity to left-vs-right sweeps, perhaps due to the asymmetric nature of the sensors in the device head. We introduced a skew factor that widened the relative probability of hits in mine vs. clutter for leftward sweeps while narrowing it for rightward sweeps, thus leaving the overall probabilities constant. Figure 3 shows that the model was able to increase its detection ability as that skew factor increased from 0 to 5, 10 and 20% probability. Again, we simply represented the direction of sweep with the signal in each decision instance, and the model inferred the rest on its own.

Finally, we combined with the signal location information whose distribution differed according to the nature of the underlying object. Mine signals were located over the center of the target while clutter signals followed a bimodal distribution on each side of the center, at a distance controlled by a separation parameter. Figure 4 shows the substantial improvement in discriminability as a function of that parameter. Again, the model was able to extract information from that continuous, probabilistic dimension to improve its detection performance without being explicitly instructed to factor it in its decision.

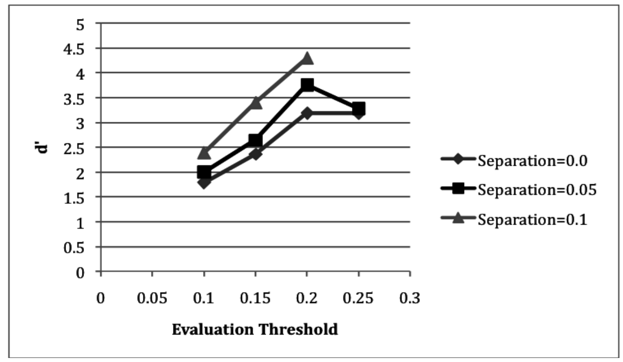

Figure 4: d' as a function of separation parameter for bimodal clutter distribution vs. central mine distribution.

Macroeconomic Decision-Making

We are currently in the process of collecting human performance data from a domain expert to calibrate, and validate and test the model. We decided to make a virtue out of necessity by generalizing our model to another situation for which data were available. In addition to providing us with validation data, this approach had the advantage of forcing us to consider our model in its full generality, not as a model of this specific task but instead of the general cognitive capacity of frequency-based decision-as making. This methodology has a number of advantages: 1) it removes degrees of freedom often used to maximize the degree of fit in task-specific models; 2) it leads to models that are applicable across a broad range of tasks, thus promoting model reuse; and 3) it provides a more stringent test of the cognitive architecture by increasing the range of constraints it has to meet.

This approach is possible in our case because of the ubiquity of frequency-based decision-making in a range of very different domains. As another instance of this domain, Blinder and Morgan (2000) designed a pair of experiments to contrast individual and group decision-making. Both experiments were framed at different levels of abstraction in terms of the macroeconomic process of controlling interest rates. They describe quite nicely the connection between macroeconomic control and sequential decision-making as:

"After the arrival of each new piece of data … policymakers ask themselves whether to adjust policy now or wait for more information -- which is precisely what our student subjects had to do."

In the first and most abstract experiment, a series of colored balls (blue or red) are drawn with replacement from an urn. The task is to determine from the time series of binary samples when and, more importantly, in which direction (red or blue) the composition of the urn switches from an even (50-50) to an uneven (70-30) distribution, a process intended to simulate the detection of an underlying change in a macroeconomic system from a series of economic indicators. Task incentives are weighted to reward early detection, but strongly punish erroneous decisions. Specifically, subjects, starting with 40 points, are awarded an extra 60 points for a correct guess but penalized one point for each sample after (or, for that matter, before) the distribution switch. Subjects were allowed an initial free experimentation period, and then played 90 rounds of the game. The subjects alternated between individual and group decision-making conditions (as indicated on some graphs) but for our purposes we will ignore those and focus on the overall significant trends. Figure 5 displays a clear increase in lag (the difference between the trial at which the decision is made and the trial at which the distribution switch happened) as a function of round (decision-making trial), averaged over 100 subjects:

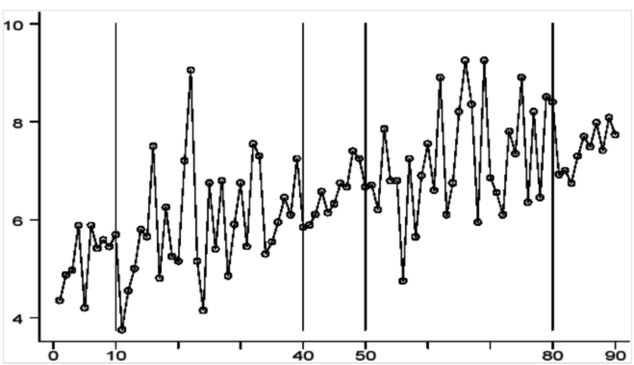

Figure 5: Increasing Average Lag with Round.

Initially subjects made a decision about 4 samples after the underlying distribution had shifted and by the end of the experiment they waited almost 8 trials. Figure 6 displays a similar increase in the percentage of correct decisions for each round. The percentage of correct decisions (about the direction of the statistical shift) increased from about 75% to about 90% after 90 rounds of practice.[1] An important question is whether the improvement in decision accuracy is solely or mostly due to the increased delay in making the decision. As we will see, different ACT-R models give differing answers to that question. Blinder & Morgan (2000) proposed three regression models to explain these empirical findings. However, they conceded that "none of these simple, intuitive models of group decision-making gets us very far."

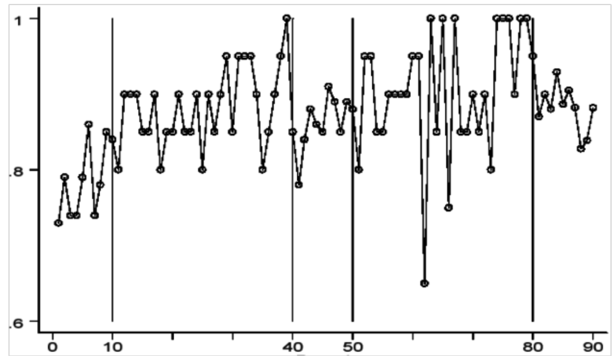

Figure 6: Increasing Percentage Correct with Round.

Procedural and Declarative Models

Lebiere & Shang (2002) proposed a cognitive model of the task based on the ability of ACT-R's procedural module to select production rules as a function of past utility, as previously applied in the economic domain by Bracht, Lebiere & Wallach (1998). The utility of a production rule is learned as a result of its actions according to the equation [2]:

$$U_i(n) = U_i(n-1) + \alpha\left[R_i(n) - U_i(n-1)\right]$$

where Ui(n) is the utility of production rule i after its nth application, Ri(n) is the reward received for its nth application, and α is the learning rate. This is a typical reinforcement learning equation, also similar to the basic Rescorla-Wagner learning rule (Rescorla & Wagner, 1972).

The key issues in applying this approach are determining (a) what the problem representation that the production rules can match is and (b) the content of the production rules. The basis for making a decision at any given point is the set of sampling outcomes up to that point, i.e. a potentially large set of binary events (red or blue balls) that would likely strain the capacity of the architectural buffers, (especially the goal and problem buffers) to which the production module has access. One representational solution would be to attempt to learn the sequence of events and use ACT-R's ability to represent such sequences as a structured set of chunks (e.g. Lebiere & Wallach, 2001) but there would be significant difficulties

[1] The increasing variability in the 10-40 and 50-80 intervals results from the switch to group decision-making, and the resulting lower sampling rate (20 groups of 5 instead of 100 individuals). Again, we will focus here on the individual performance intervals.

[2] Lebiere & Shang's original model used an older form of that equation that was standard in ACT-R at the time. The effect of the two equations is substantially similar, and the new equation works better with continuous reward, such as in this case.

both in remembering the sequence of events on the fly and in using a complex, abstract representation of it in making a binary blue-vs-red decision. An alternative is to represent the sequence in a more compact form while preserving the bulk of its information. The solution finally adopted was to keep a running count of the difference between red and blue balls, and use that difference as the basis for making a decision.

While that decision is in keeping with the frequency basis of the problem, it has substantial shortcomings, specifically that the ratio of that difference to total samples is the most relevant information in estimating probabilities, and more problematically that the early results should be discounted because they likely happened before the distribution shift and thus are likely to reflect random sampling fluctuations rather than contain information about the shift. However, there is no closed form solution for such discounting, and so we adopted the difference representation for tractability reasons. The upside of this decision is that it made the production rules that can apply to that representation relatively straightforward. Specifically, the production set is listed in Table 1. The production set consists of a set of "take" rules that make a decision (picking the color with the greater number of samples) when the difference between the number of red and blue samples reaches a specific threshold (from 3 to 6 or more), plus a "pass" rule that requests another sample and updates the count difference based on the outcome.

Take=3	Decide greater of red or blue if difference = 3
Take=4	Decide greater of red or blue if difference = 4
Take=5	Decide greater of red or blue if difference = 5
Take>=6	Decide greater of red or blue if difference >= 6
Pass	Request another draw and update difference

Table 1: Production Set for Procedural Model.

When the difference in counts is two or fewer (such as at the start of a round), only the Pass rule matches. It is always selected and requests and processes another draw. When the difference reaches a threshold covered by one of the Take rules, the system now has a choice between that rule and the Pass rule, which always applies. The production rule with the highest utility is selected according to a softmax (a.k.a. Boltzmann) procedure, which is equivalent to adding noise to the utilities and selecting the highest. If the Take rule is selected, a reward is received from the system and is used to update the utility of the production rules that have been selected up to this point, which consists of a series of Pass rules and the final Take rule. The reward received by each instantiation of the Pass rule is discounted by the amount of time between that application of the rule and the reward. The process then repeats for the given number of rounds. The model accounted for the main results quite nicely, displaying a similar increase in accuracy and lag as human subjects.

We applied the declarative model of landmine detection to the balls-and-urn paradigm to test its generalization and contrast it with the procedural model. The model generalization was relatively straightforward, as the only thing to specify was the nature of events (replacing abstract labels of hits and misses with red and blue balls) and defining the ground truth for learning purposes (in this case, the direction of the shift in distribution). Space limitations prohibits us from reporting detailed results, but the model displayed an increase in accuracy from about 70% to about 90%, similar to the subject data. However, this improvement did not arise from an increase in selectivity (and associated lag) but instead an improvement in its estimation accuracy. It raises the question as to whether and how the model could learn to refine its decision threshold. While the threshold is currently fixed, the model could represent it explicitly and adjust it up with failure and down with successes based on its evaluations, a key feature of the procedural model.

Conversely, we computed the maximum d' on the landmine detection task for various Take productions of the procedural model. Productions corresponding to a difference threshold of 3, 4, 5 and 6 respectively yielded d' of 3.6, 4.2, 4.6 and 4.9 respectively. That excellent performance is marred by the fact that real-world data is not likely to feature the wide gap in hit probabilities between mine and clutter used here, but will instead emphasize the discovery of subtle statistical regularities that the procedural model cannot accomplish because of its hardwired representation. Moreover, generalization to continuous systems, such as Blinder and Morgan's second experiment which features real-valued quantities such as interest rates and inflation rates and for which subject performance is very similar to the discrete balls experiment, would also be highly problematic.

Thus, the two models can also be seen as fundamentally complementary strategies. The procedural model focuses on the adjustment of the selectivity threshold, which the declarative model lacks. However, the procedural model is unable to learn the most effective representation of context, and in particular cannot detect subtle statistical relationships between input dimensions and decision effectiveness. The declarative model focuses on the representation of context from statistical patterns, which the procedural model lacks. While it currently uses a constant decision threshold, it could represent that threshold explicitly and adjust it adaptively, bumping it down when making a correct decision and bringing it back up in case of failure. The declarative evaluation process, with its transformation of context (evaluation) and signal into a revised evaluation, can even be seen as an explicit version of the Rescorla-Wagner rule. Thus exploring both domain space and model space might ultimately lead us to both broader explanatory coverage and architectural unification.

Conclusion

This effort to generalize across models of frequency-based decision making within a cognitive architecture reveals the ability of both a procedural and a declarative model to explain the Blinder and Morgan (2000) results. These indeterminant results could be considered a failure, if their implications regarding human strategic variability are ignored. Doing so, however, is likely to hinder efforts to explain skilled performance in terms of cognitive mechanisms and further understand their neural substrate. Behavioral studies show that glossing over strategic variability muddles understanding of thought processes (Siegler, 1987). Recent work examining the effects of strategic variability within a cognitive architecture, show how different neural patterns implement different strategies for complex arithmetic calculation (Rosenberg-Lee, Lovett, & Anderson, 2009) further supporting this conclusion. Thus, our pursuit of theoretical breadth and parsimony has unexpectedly produced results that highlight an important feature of human thought: intelligent systems' flexibility to devise and use a multiple problem solution strategies. Analogous to the way that genetic variability and genetic change can influence adaptation, so too can strategic variability and change. Understanding strategic variability and its sources at the level of cognitive mechanisms appears necessary to model complex cognition accurately within cognitive architectures and to advance understanding the neural foundations of learning and cognitive performance.

Acknowledgments

This research was funded by Subaward #162-10 from the Leonard Wood Institute.

References

Anderson, J. R. 2007. How Can the Human Mind Occur in the Physical Universe? New York, NY: Oxford University Press.

Anderson, J. R., and Lebiere, C. 1998. The Atomic Components of Thought. Mahwah, NJ: Lawrence Erlbaum Associates.

Blinder, A. S., and Morgan, J. 2000. Are Two Heads Better Than One? An Experimental Analysis of Group vs. Individual Decision-making. In *Working Paper Series* No. 7909, National Bureau of Economic Research.

Bracht, J., Lebiere, C., and Wallach, D. 1998. On the Need of Cognitive Game Theory: ACT-R in Experimental Games with Unique Mixed Strategy Equilibria. In *Joint Meetings of the Public Choice Society and the Economic Science Association.* New Orleans, LA.: Joint Meetings of the Public Choice Society and the Economic Science Association.

Gonzalez, C., and Lebiere, C. 2005. Instance-Based Cognitive Models of Decision Making. In *Zizzo, D., and Courakis, A. eds. Transfer of Knowledge in Economic Decision Making.* New York, NY.: Palgrave McMillan.

Lebiere, C., and Wallach, D. 2001. Sequence Learning in the ACT-R Cognitive Architecture: Empirical Analysis of a Hybrid Model. In *Sun, R., and Giles, L. eds. Sequence Learning: Paradigms, Algorithms, and Applications.* Germany.: Springer LNCS/LNAI.

Lebiere, C., and Shang, J. 2002. Modeling Group Decision Making in the ACT-R Cognitive Architecture. In *Proceedings of the 2002 Computational Social and Organizational Science (CASOS) Conference*. June 21-23, Pittsburgh, PA.

Lebiere, C., Gonzalez, C., and Warwick, W. 2009. A Comparative Approach to Understanding General Intelligence: Predicting Cognitive Performance in an Open-Ended Dynamic Task. In *Proceedings of the Second Artificial General Intelligence Conference (AGI-09)*. Amsterdam, Paris.: Atlantis Press.

Lipkus, I. M. 2007. Numeric, Verbal, and Visual Formats of Conveying Health Risks: Suggested Best Practices and Future Recommendations. *Medical Decision Making* 27(5): 696-713.

Rescorla, R. A., and Wagner, A. R. 1972. A Theory of Pavlovian Conditioning: Variations on the Effectiveness of Reinforcement and Nonreinforcement. In *Black, A. H., and Prokasy, W. F. eds. Classical Conditioning: II. Current Research and Theory,* 64-99. New York, NY.: Appleton-Century-Crofts.

Rosenberg-Lee, M., Lovett, M. C., and Anderson, J. R. 2009. Neural Correlates of Arithmetic Calculation Strategies. *Cognitive Affective, and Behavioral Neuroscience* 9(3): 270-285.

Sanner, S., Anderson, J. R., Lebiere, C., and Lovett, M. C. 2000. Achieving Efficient and Cognitively Plausible Learning in Backgammon. In *Proceedings of The Seventeenth International Conference on Machine Learning*. San Francisco, Calif.: Morgan Kaufmann.

Sedlmeier, P., and Betsch, T. 2002. Etc. Frequency Processing and Cognition. Oxford, UK.: Oxford University Press.

Siegler, R. S. 1987. The Perils of Averaging Data Over Strategies: An example from Children's Addition. *Journal of Experimental Psychology: General* 116: 250-264.

Staszewski, J. (2004). Models of Expertise as Blueprints for Cognitive Engineering: Applications to Landmine Detection. Proceedings of the 48th Annual Meeting of the Human Factors and Ergonomics Society, New Orleans, LA, 20-24 September.

Stewart, T. C., West, R., and Lebiere, C. 2009. Applying Cognitive Architectures to Decision Making: How Cognitive Theory and The Equivalence Measure Triumphed in the Technion Prediction Tournament. To *appear in Proceedings of the 2009 Cognitive Science Conference*.

Taatgen, N., Lebiere, C. and Anderson, J.R. 2006. Modeling Paradigms in ACT-R. In *Sun, R. (Ed) Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation.* New York, NY: Cambridge University Press.