The Boundaries of Instance-Based Learning Theory for Explaining Decisions from Experience

Cleotilde Gonzalez

Dynamic Decision Making Laboratory

Social and Decision Sciences Department

Carnegie Mellon University

coty@cmu.edu

208 Porter Hall, Pittsburgh, PA 15213

Phone: 412-268-6242

Fax: 412-268-6938

Acknowledgements: This is a summary of research from many years, and thus I owe recognition to a large number of organizations that supported my work, including the Army Research Laboratory, the Office of Naval Research, the National Science Foundation, and the Army Research Office, among others. I also owe recognition to doctoral students, post-doctoral fellows and other research staff in the dynamic decision making laboratory.

Abstract

Most demonstrations of how people make decisions in risky situations rely on *decisions from description*, where outcomes and their probabilities are explicitly stated. But recently, more attention has been given to *decisions from experience* where people discover these outcomes and probabilities through exploration of the problems. More importantly, risky behavior depends on how decisions are made (from description or experience), and although Prospect Theory explains decisions from description, a comprehensive model of decisions from experience is yet to be found. Instance-Based Learning Theory (IBLT) explains how decisions are made from experience through interactions with dynamic environments (Gonzalez, Lerch, & Lebiere, 2003). The theory has shown robust explanations of behavior across multiple tasks and contexts, but it is becoming unclear what the theory is able to explain and what it does not. The goal of this chapter is to start addressing this problem. I will introduce IBLT and a recent cognitive model based on this theory: the IBL model of repeated binary choice; then I will discuss the phenomena that the IBL model explains and those that the model does not. The argument is for the theory's robustness but also for clarity in terms of concrete effects that the theory can or cannot account for.

The Boundaries of Instance-Based Learning Theory to Explaining Decisions from Experience

Theories that explain human decision making have traditionally involved principles and developments from Economics and Psychology, and for many years these two disciplines have proposed what appear as conflicting mechanisms and explanations. On the one hand, Economists have assumed humans to be utility maximizers (i.e., "rational"), while Psychologists aimed at demonstrating the many different decision situations in which humans are not utility maximizers (i.e., "irrational"). A major breakthrough in behavioral decision research was the shift of attention from particular examples that dispute expected utility theory to explanations of *how* people make decisions through Prospect Theory (Kahneman & Tversky, 1979). This theory has been a prominent model used to explain and generalize deviations from expected utility theory.

While demonstrating the explanatory power of Prospect Theory, researchers have traditionally used monetary gambles (i.e., "prospects") that explicitly state outcomes and associated probabilities. People are presented with a description of the alternatives and they are asked to make a choice based on the conditions described, they are asked to make *decisions from description*. For example:

Which of the following would you prefer?

A: a .8 chance to get $4 and .2 chance to get $0

B: get $3 for sure

Using decisions from description, researchers have investigated a large number of situations in which people behave against utility maximization and in agreement with Prospect Theory, producing an impressive list of “heuristics and biases” (Tversky & Kahneman, 1974). Through the years, these consistent deviations from rational behavior have been identified, replicated, and extended upon using laboratory experiments, to the point where this type of research has dominated the field of behavioral decision making for the past six decades.

However, despite the many years of effort, we have only limited answers to the question of *how* people make decisions; rather, most research has aimed at demonstrating how people don't make decisions. The large collection of cognitive biases cannot be all explained by one comprehensive theory and most importantly, we do not know how the biases develop and how do they emerge in the first place. As a result, we know little of how to prevent them. Most empirical studies up to date focus on the observable processes such as choice selection, and ignore cognitive processes that lead to choice, such as recognizing alternatives, deciding when to search for information, evaluating and integrating possible outcomes, and learning from good and bad decisions, among other processes.

A recent development in decision sciences has great potential to expand our understanding and provide insights into the decision making process. A shift of attention to how decisions are made from experience (i.e., *decisions from experience*), rather than from explicit description of options, opens a window towards a better understanding of cognitive processes that including: information search, recognition and similarity processes, integration and accumulation of information, feedback, and learning. Researchers use experimental paradigms that involve repeated decisions rather than one-shot decisions, the estimation of possible outcomes and probabilities based on the observed outcomes rather than from a written description, and learning from feedback. All of which are natural processes for making decisions in many real-world situations in which alternatives, outcomes, and probabilities are unknown. The experimental paradigm often involves two alternatives, represented by two unlabeled buttons, each representing a probability distribution of outcomes that is unknown to participants. Clicking a button yields an outcome as a result of a random draw from the alternative's distribution. Although there are multiple paradigms for the study of decisions from experience (Hertwig & Erev, 2009; Gonzalez & Dutt, 2011), a common paradigm is the "sampling" paradigm (see Figure 1), in which people are able to explore the outcomes of the options without real consequences before they decide to make a final choice.

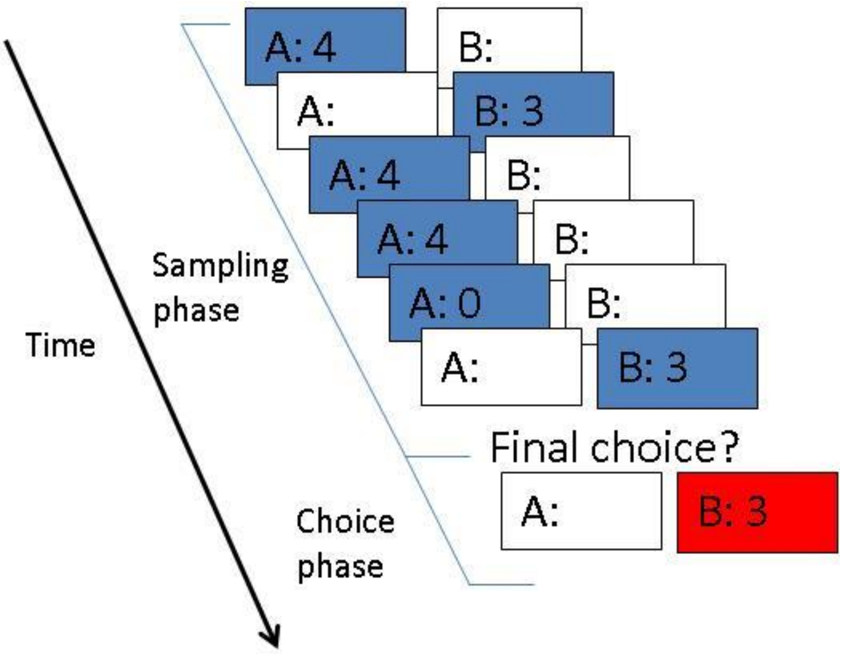

Figure 1. The sampling paradigm of decisions from experience.

A key observation that contributed to the initial success of the theoretical development of decisions from experience was the "description-experience gap" (Hertwig, Barron, Weber, & Erev, 2004): that the choice that an individual makes depends on how information about the problem is acquired (from description or experience); particularly in problems involving outcomes with low probabilities

(probabilities less than .2, "rare events"). A robust finding across a range of paradigms for decisions from experience is that people behave as if rare events have *less* impact than they deserve according to their objective probabilities. More importantly, this finding contradicts the prediction from prospect theory that people behave as if rare events have *more* impact than they deserve. However, this theory only applies to "simple prospects with monetary and stated probabilities" (Kahneman & Tversky, 1979 pp. 274). Thus, although prospect theory seems to provide good explanations for decisions from description, findings from decisions from experience may contradict those predictions from prospect theory in many cases (Hertwig, 2012).

Although prospect theory (Kahneman & Tversky, 1979) has been a prominent model to explain human-choice behavior in descriptive choices, a comprehensive model that can explain decisions from experience has not yet been found. In fact, a challenge in understanding the cognitive processes involved in making decisions from experience is the proliferation of highly task-specific cognitive models that often predict behavior in a particular task, but fail to also explain behavior even in closely related tasks (see discussions in Gonzalez & Dutt, 2011; Lejarraga, Dutt, & Gonzalez, 2012). Gonzalez and colleagues have attempted to address this challenge by providing multiple demonstrations of how cognitive computational models based on one theory, Instance-Based Learning Theory (IBLT; Gonzalez et al., 2003), account for human behavior in a large diversity of tasks where decisions are made from experience. Recently, they have demonstrated that the same computational model based on IBLT, without modifications, is able to account for multiple variations of the dual choice paradigms commonly used to study decisions from experience (e.g., Gonzalez & Dutt, 2011; Lejarraga et al., 2012).

In what follows, I summarize IBLT as a general theory of decision making in dynamic tasks. I discuss how IBLT has accounted for decision making behavior on a wide range of tasks that vary in their dynamic characteristics across a taxonomy of dynamic tasks. I then concentrate on a model proposed for the study of decisions from experience in the least dynamic task of the taxonomy, the repeated choice paradigms (e.g., Figure 1). Next I present a set of phenomena in decision sciences that the IBL model has shown to explain and predict accurately. I will also summarize the type of learning and decisions from experience phenomena that the IBL model in its current form does not explain, and conclude on some ideas and plans to expand the current IBL model.

Instance-Based Learning Theory

Instance-Based Learning Theory (IBLT) was developed to explain human decision making behavior in dynamic tasks (Gonzalez et al., 2003). In dynamic tasks, individuals make repeated decisions attempting to maximize gains over the long run (Edwards, 1961; 1962; Rapoport, 1975). According to Edwards (1962), dynamic decision tasks are characterized by decision conditions that change spontaneously and with time, inaction, and as a result of previous decisions.

Based on evidence from studies in naturalistic environments (Dreyfus & Dreyfus, 1986; Klein, Orasanu, Calderwood, & Zsambok, 1993; Pew & Mavor, 1998; Zsambok & Klein, 1997), laboratory studies with dynamic computer simulations (Microworlds) (Brehmer, 1990, 1992; Gonzalez, 2004, 2005; Kerstholt & Raaijmakers, 1997), theoretical studies of decisions under uncertainty (Gilboa & Schmeidler, 1995, 2000), and other theories of learning in dynamic decision making (Dienes & Fahey, 1995; Gibson, Fichman, & Plaut, 1997); IBLT proposed that decisions in dynamic tasks were made possible by referencing experiences from past similar situations, and applying the decisions that worked in the past. IBLT's most important development was the description of the learning process and mechanisms by which experiences may be built, retrieved, evaluated, and reinforced during the interaction with a dynamic environment.

IBLT characterizes learning in dynamic tasks by storing "instances" in memory as a result of having experienced decision making events. These instances are representations of three elements: a situation (S), which is defined by a set of attributes or cues; a decision (D), which corresponds to the action taken in situation S; and a utility or value (U), which is expected or received for making a decision D in situation S. IBLT proposes a generic decision making process through which SDU instances are built, retrieved, evaluated, and reinforced (see detailed description of this process in Gonzalez et al., 2003); with the steps consisting of: recognition (similarity-based retrieval of past instances), judgment (evaluation of the expected utility of a decision in a situation through experience or heuristics), choice (decision on when to stop information search and select the optimal current alternative), execution (implementation of the decision selected), and feedback (update of the utility of decision instances according to feedback). The decision process of IBLT is determined by a set of learning mechanisms needed at different stages, including: Blending (the aggregated weighted value of alternatives involving the instance's utility weighted by its probability of retrieval), Necessity (the decision to continue or stop exploration of the environment), and Feedback (the selection of instances to be reinforced and the proportion by which the utility of these instances is reinforced).

IBLT and IBL Models

To test theories of human behavior, we use *computational models*: representations of some or all aspects of a theory as it applies to a particular task or context. Thus, the value of models is that they can solve concrete problems and provide explicit mathematical and computational representations of a theory, which can then be used to make predictions of behavior.

IBLT constructs and processes were implemented into a computational model (called Cog-IBLT) that helped make the theory more explicit, transparent, and precise (Gonzalez et al., 2003). Cog-IBLT demonstrated the overall mechanisms and learning process proposed by the theory in a dynamic and complex resource allocation task (the "water purification plant", reported in Gonzalez et al., 2003). Cog-IBLT was constructed within the ACT-R cognitive architecture (Anderson & Lebiere, 1998), using the cognitive mechanisms existent in ACT-R. Specifically, Cog-IBLT used the ACT-R's experimentally-derived mathematical representations of: *Activation* (a value that determines the usefulness of an instance from memory and experience and the relevance of the instance to the current context); *Partial Matching* (a value that determines the similarity of instances and the retrieval of instances that may be only similar to a current environmental situation); and *Retrieval Probability* (a value representing the probability of retrieving an instance as a function of Activation and Partial Matching). This model also used a modified version of the concept of *Blending* proposed in Lebiere's dissertation (2008): An aggregate or combination of values of multiple instances in memory. Through a series of "simulation experiments," the Cog-IBLT demonstrated the explanatory and predictive potential of IBLT, as it closely approximated the learning process from human data in the water purification plant task.

As a general theory of dynamic decision making (DDM), IBLT aims at addressing a wide range of dynamic tasks. Edwards (1962) proposed an initial taxonomy of dynamic tasks, ranging from the least dynamic, where actions are sequential in an environment that is constant and where neither the environment nor the individual's information about the environment is affected by previous decisions (as the repeated choice task in Figure 1); to the most dynamic, where the environment and the individual's information about it changes over time and as a function of previous decisions (as in the water purification plant task used in Cog-IBLT). This taxonomy was later extended to include an even more dynamic characteristic in Edwards' taxonomy: that decisions are made in real time, and thus their outcomes depend on the time at which the decision is made (Brehmer, 1992; Hogarth, 1981).

After Cog-IBLT, many IBL models have been developed in a wide variety of dynamic decision making tasks across the taxonomy of dynamic tasks from the most dynamic to the least dynamic task, including: dynamically-complex tasks (Gonzalez & Lebiere, 2005; Martin, Gonzalez, & Lebiere, 2004), training paradigms of simple and complex tasks (Gonzalez, Best, Healy, Kole, & Bourne, 2010; Gonzalez & Dutt, 2010), simple stimulus-response practice and skill acquisition tasks (Dutt, Yamaguchi, Gonzalez, & Proctor, 2009), and repeated binary-choice tasks (Lebiere, Gonzalez, & Martin, 2007; Lejarraga et al., 2012) among others.

A recent IBL model has shown generalization across multiple tasks that share structural similarity with the paradigms used to study decisions from experience (as in Figure 1). Although these tasks are the least dynamic in the taxonomy of Edwards (1962), they shown great potential to develop and test IBLT, given their simplicity. An IBL model was initially built to predict performance in individual repeated binary-choice tasks. Motivated by the work of Erev and Barron (2005), we built a model of repeated binary choice based on IBLT but within the ACT-R architecture (Lebiere et al., 2007). Erev and Barron (2005) demonstrated robust deviations from maximization in repeated binary choice and proposed the Reinforcement Learning Among Cognitive Strategies (RELACS) model, which closely captures human data and outperforms other models. We argued for a simpler model, the IBL model, which was able to fit the data as well as RELACS (Lebiere et al., 2007).

The IBL model's development took an important turn when it was submitted to the Technion Prediction Tournament (TPT; Erev, Ert, Roth et al., 2010), a modeling competition that involved fitting and prediction phases, where the model authors were given a data set to fit their models to and were evaluated in a novel data set. The IBL model was developed independently and outside from ACT-R, and the mechanisms of this model were isolated from all the other ACT-R mechanisms (see Gonzalez, Dutt, & Lebiere, in press for a validation of this model within ACT-R and outside of ACT-R). Although this model did not win the TPT, the model's transparency, simplicity, and flexibility outside of ACT-R have been an advantage to recent developments. The IBL model has now been shown to predict performance better than the winner models of the TPT (Gonzalez & Dutt, 2011; Lejarraga et. al., 2012); to predict performance in a variety of repeated binary-choice tasks, probability-learning tasks, and dynamic choice task across the multiple paradigms of decisions from experience; and at the individual and team levels (Gonzalez & Dutt, 2011; Gonzalez, Dutt, & Lejarraga, 2011; Lejarraga et al., 2012). The discussions from this point on will refer to this particular IBL model, which is explained in detail next.

The IBL Model of Decisions from Experience

Instances in a model of the decision from experience paradigms (e.g., that shown in Figure 1) have a much simpler representation compared to instances in Cog-IBLT or in other IBL models. The instance structure is simple because the task structure is also simple. Each instance consists of a label that identifies a decision option in the task and the outcome obtained. For example, (Left, $4) is an instance where the decision was to click the button on the left side and the outcome obtained was $4. The details of this IBL model and its relevance were fully explained in Gonzalez and Dutt (2011), but the main aspects of this model are summarized here.

The IBL model of decisions from experience ("IBL model" hereafter) assumes that choices from experience are based on either a repetition of past choices (i.e., "inertia") or on the aggregation of past experiences (i.e., “instances”) of payoffs in memory that have been observed as a result of past choices (i.e., "blending"). At trial $t = 1$, the model starts with a random choice between the two options. Then, in each trial $t > 1$, the model first applies a probabilistic rule (based upon a free parameter called *pInertia*) to determine whether to repeat its choice from the previous trial or not. If this probabilistic rule fails, then inertia does not determine the choice and the model chooses the option with the highest *blended* value. An option's blended value is a weighted average of all observed payoffs on that option in previous trials. These observed payoffs are stored as instances in memory and are weighted such that payoffs observed more frequently and recently receive a higher weight compared to the less frequent and distant payoffs. This weight is a function of the recency and frequency of the instances’ use, where the instance contains the observed payoffs. Formally, the model works as follows:

In $t = 1$ choose randomly between the two choice options (1)

For each trial $t > 1$,

If the draw of a random value in the uniform distribution *U* (0, 1) < *pInertia*,

Then

Repeat the choice as made in the previous trial

Else

Select an option with the highest blended value as per Equation 2 (below)

The blended value *V* of option *j* is:

$$V_j = \sum_{i=1}^{n} p_{ij} x_{ij} \tag{2}$$

where x_{ij} is the observed payoff in instance i for the option j, and p_{ij} is the probability of retrieving that instance for blending from memory (Gonzalez & Dutt, 2011; Lejarraga et al., 2012). Since the sampling paradigm involves a binary-choice with two options, the values of j can be either 1 or 2 (i.e., right or left choice options). Thus, the blended value of an option j is the sum of all x_{ij} stored in instances in memory, weighted by their probability of retrieval p_{ij}. The *n* value is the number of different instances containing observed payoffs on option *j* up to the last trial. For example, if by trial *t = 2*, option *j* revealed 2 different payoffs stored in two instances, then *n* = 2 for option *j*. If the two observed payoffs on option *j* are the same in the previous two trials, then only one instance is created in memory and *n* = 1.

In any trial, the probability of retrieving from memory an instance *i* containing a payoff observed for option *j* is a function of that instance's activation relative to the activation of all other instances that contain observed payoffs l occurring within the same option. This probability is given by:

$$p_{ij} = \frac{e^{\frac{A_i}{\tau}}}{\sum_l e^{\frac{A_l}{\tau}}} \tag{3}$$

where l refers to the total number of payoffs observed for option j up to the last trial, and τ is a noise value defined as $\sigma \cdot \sqrt{2}$ (Lebiere, 1998). The σ variable is a free noise parameter expected to capture the imprecision of recalling instances from memory from one trial to the next.

The activation of each instance in memory depends upon the activation mechanism originally proposed in the ACT-R architecture (Anderson & Lebiere, 1998). The IBL model uses a simplified version of that activation mechanism. In each trial t, activation A of an instance i is

$$A_i = \ln\left[\sum_{t_i \in \{1,\ldots,t-1\}} (t - t_i)^{-d}\right] + \sigma \cdot \ln\left(\frac{1-\gamma_i}{\gamma_i}\right) \tag{4}$$

where d is a free decay parameter, and t_i refers to previous trials when payoff contained in the instance i was observed (if a payoff occurs for the first time in a trial, a new instance containing this payoff is created in memory). The summation will include a number of terms that coincides with the number of times that a payoff has been observed after it was created (the time of creation of instance itself is the first timestamp). Therefore, an instance's activation containing a payoff increases with the frequency of observing that payoff (i.e., by increasing the number of terms in the summation) and with the recency of observing that payoff (i.e., by small differences in $t - t_i$). The decay parameter d affects the activation of the instances directly, as it captures the rate of forgetting. The higher the value of the d parameter, the faster the decay of instances' activations in memory is.

The γ_i term is a random draw from a uniform distribution defined between 0 and 1, and $\sigma \cdot \ln\left(\frac{1-\gamma_i}{\gamma_i}\right)$ represents the Gaussian noise that is important for capturing variability in behavior from one trial to the next. The σ variable is the same noise parameter defined in equation 3 above. A high σ implies a high noise in activation.

The most recent developments of the IBL model of decisions from experience are important given the simplicity of this model and the broad predictions that it can make (e.g., Gonzalez & Dutt, 2011; Gonzalez et al., 2011; Lejarraga et al., 2012). Next section describes some examples of what the model is able to explain and what the model in its current form does not explain. All examples below rely on two parameters: the decay, d, and the noise, σ with values 5.0 and 1.5 respectively. However, the models reported below vary in the inclusion or not of the *pInertia* parameter (see Dutt & Gonzalez, 2012 for a discussion on the value of this parameter), and also on the specific values of the parameters. As explained next, we have used a fit and generalization procedure, in which the parameters values are fit to particular data sets and then used these parameters to predict the behavior in a new data set.

What the IBL model explains and what it does not explain

Existent demonstrations from IBL models suggest the generality of the theory, and not only the descriptive power of the theory but the explanatory one. That is, the theory not only describes the kind of constructs and processes existent in dynamic decision making, but it helps explain why decision making in dynamic tasks occur in the way described. But with generality and robustness also comes the lack of specificity: What are the effects and phenomena that the IBL model can explain and predict? Here we first summarize this tradeoff between generality and specificity, then we present the concrete phenomena that the model in its current form is capable and not capable of explaining.

What the IBL model explains.

Two comprehensive and important demonstrations of the IBL model's robustness are the fitting and predictions obtained against a large and publicly available data set, the TPT (Erev et al., 2010). TPT was a competition in which different models were submitted to predict choices made by experimental participants. Competing models were evaluated following the generalization criterion method (Busemeyer & Wang, 2000): they were fitted to choices made by participants in 60 problems (the estimation set) and later tested using the parameters that best fitted the estimation data set to predict a new set of choices in 60 problems (the test set). This process of fitting and generalization procedure is useful as generalization is regarded as pure prediction of behavior.

TPT involved 2 types of experimental paradigms of decisions from experience, Sampling and Repeated choice; and all the problems in the TPT involved a choice between two options:

Safe: M with certainty

Risky: H with probability Ph; L otherwise (with probability 1-Ph)

A safe option offered a medium (M) payoff with certainty, and a risky option that offered a high (H) payoff with some probability (pH) and a low (L) payoff with the complementary probability. M, H, pH, and L were generated randomly, and a selection algorithm assured that the 60 problems in each set differed in domain (positive, negative, and mixed payoffs) and probability (high, medium, and low pH).

An example of the IBL model's predictions has been reported by Lejarraga et al. (2012) and reproduced in Figure 2. Figure 2 shows the learning curves on the proportion of risky choices (P-Risky) of each of the 60 problems in the test set. As can be seen, the IBL model accurately predicted learning in most of the problems (see detailed tests in Lejarraga et al., 2012).

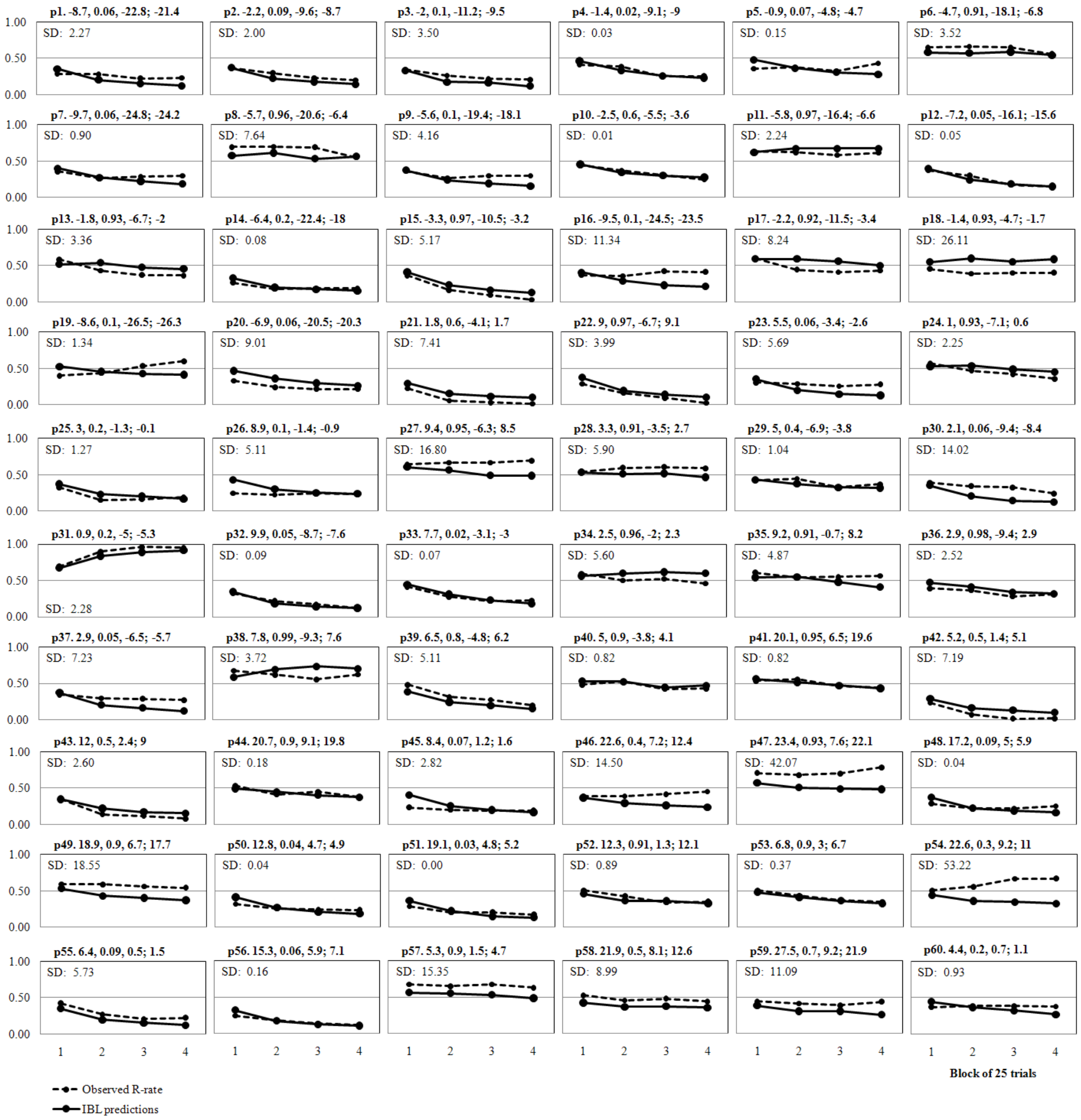

Figure 2. Learning curves from human and IBL model data in the test set of the TPT. Each panel represents one of the 60 problems, each problem ran for 100 trials (both for the IBL model and human data),and the panels show the proportion of risky choices averaged in blocks of 25 trials. The SD in each graph denotes the squared distance between the observed R-rate and the IBL predictions across 100 trials. The IBL model was run in exactly the same experimental paradigm as humans were. The model included the same simulated participants as the human data set.

The 60 problems represent a large diversity of behavioral effects, and in creating this diversity of problems, the organizers of the TPT (Erev et al., 2010) aimed at extending the traditional view of using counter-examples of particular behavioral effects by demonstrating the robustness of general learning effects. This demonstration and additional ones in Lejarraga et al. (2010) and in Gonzalez and Dutt (2011) indicate the IBL model's ability to capture these general learning effects too.

However, reliance on quantitative model comparison and numerical model predictions may lead this work to need of a "help line" (Erev et al., 2010) to guide potential users on what phenomena that this model can explain and the predictions that it can and cannot currently make. Although the TPT problems represent a large diversity of behavioral effects, these are difficult to isolate. This is because the problems were created with an algorithm that randomly selected outcomes and probabilities in such a way that 1/3 of the problems involve rare High outcomes (Ph<0.1) and about 1/3 involve rare Low outcomes (Ph>0.9); also 1/3 of the problems are in the gain domain (all outcomes are positive) and 1/3 are in the loss domain (all outcomes are negative). Thus, effects such as those found in other studies (e.g., Erev & Barron, 2005) may be difficult to isolate in the TPT's diverse problem sets.

We aim to address the question of robustness and specificity for the IBL model in the following sections, where I summarize results from the model in data sets where different type of phenomena were clearly identified: payoff variability effect, underweighting of rare events, loss rate effect, individual differences (Erev & Barron, 2005), probability matching, and adaptation to nonstationary environments (Lejarraga et al., 2012).

The payoff variability, underweighting of rare events, and loss rate effects.

Erev and Barron (2005) demonstrated robust deviations from maximization in repeated binary choice tasks. These deviations are classified into three main effects: payoff variability, underweighting of rare events, and loss rate.

The *payoff variability* effect refers to a tendency to increase exploration when payoff variability is associated with an alternative of higher expected value (Erev & Barron, 2005). The *underweighting of rare events* effect refers to the tendency to believe that the greater value and least probable outcome is less probable than its objective probability in decisions from experience (Erev & Barron, 2005; Hertwig et al., 2004), and the *loss rate* effect indicates that people sometimes tend to prefer alternatives that minimize losses over those that maximize gains. Here we demonstrate that the same IBL model can explain all three effects in all the problems presented in Erev and Barron (2005).

A replication of Erev & Barron's payoff variability effect in three problems and IBL model predictions.

To calibrate the parameters of the IBL model, we first replicated the payoff variability effect with human participants, using the following three problems (Haruvy & Erev, 2001; Erev & Barron, 2005):

Problem 1.	H	11 points with certainty
	L	10 points with certainty
Problem 2.	H	11 points with certainty
	L	19 points with probability 0.5
		1 otherwise
Problem 3.	H	21 points with probability 0.5
		1 otherwise
	L	10 points with certainty

All three problems show a choice between a high alternative with an expected value of 11 points and a low alternative with an expected value of 10 points, but the problems differ on the variance of the two payoff distributions. We developed a computer program for data collection and we ran an experiment where each of 60 participants, undergraduate and graduate students at Carnegie Mellon University, worked on one of the three problems. We followed almost identical instructions as in the original experiments: individuals did not receive any information about the payoff structure. They were told that their task was to select one of the alternatives by clicking on one of two unmarked and masked buttons on the screen and were not informed of the trial number. They were provided with the payoff value of the button they clicked on. Payoffs were drawn from the distribution associated with the selected button. There are two differences between our methods and Erev and Barron's (2005): (1) we did not use a performance-based incentive structure. Participants were paid a flat fee for performing the repeated choice task, and (2) we ran 400, rather than 200, trials for all problems to better explore learning effects. The average proportions of maximization (i.e., Pmax, the rate choices with the highest expected value) in our data set are very similar to those reported in Erev and Barron (2005). The average Pmax for the second 100-problem block (i.e., Pmax2) was 0.82, 0.61, and 0.50 for Problems 1, 2, and 3 respectively (compared to .90, .71, and .57 in Erev and Barron (2005)). The slight but generally lower Pmax2 values in our replication may be due to the difference in the performance-based incentive.

Figure 3 shows the proportion of maximization (Pmax) choices from humans (dark lines) and those from the IBL model (dotted lines) in each of the three problems.

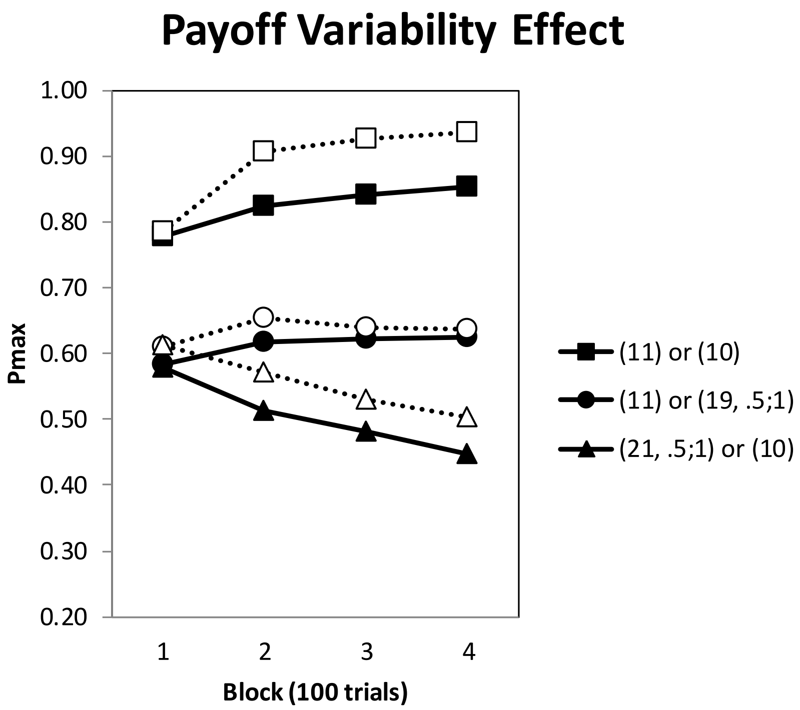

Figure 3. The payoff variability effect in Human (dark lines) and IBL Model (dotted lilnes). data. The graph shows the Proportion of maximization (Pmax) in each block of 100 trials, for a total of 400 trials. The IBL model was run in exactly the same experimental paradigm as humans were. The model included the same simulated participants as the human data set.

These learning curves illustrate that, as expected from the original experiments, an increase in payoff variability impairs maximization. Payoff variability for the high alternative decreases maximization over time. The payoff variability effect arises from the Blending mechanism (Equation 2) and the dynamics of the task values (the IBL model here does not include inertia). The model selects the option with the highest blended value; this is clear in problem 1, where the selection of the maximum option (11) is only influenced by the noise in activation (Equation 4) and in the retrieval of instances (Equation 3). In Problem 2, the model retrieves some instances of the maximum value in the risky option, 19, 50%. This makes the proportion of maximization less extreme than in problem 1, as the model would select the risky option more often because it results in the maximum Blended value. In Problem 3, the risky alternative provides some higher payoffs (e.g., 21), half of the time which raises its expected value and leads to its selection more often. But the value of the risky alternative appears to quickly even out or decrease over time as a series of poor payoffs (e.g., 1) may lower its expected value and make the certain alternative (i.e., 10) more attractive, which in turn would increase the activation of this option by its more frequent selections.

Additional demonstrations of IBL predictions of the Payoff Variability, underweighting of rare events, and loss rate effects.

We ran the IBL model in the 40 problems reported in Erev and Barron (2005), which belong to the three effects described above. We ran the IBL model in each problem over the course of 400 trials for 100 simulated participants. The set of simulations resulted in the predicted learning curves summarized as the average Pmax in four blocks of 100 trials each. Figure 4 shows the learning curves for humans and for the IBL model. The Pmax per block (100 trials in each block) is shown for each of the 40 problems from Erev and Barron (2005)[1]. The figure shows that the IBL model can account for problems that demonstrate the *payoff variability* effect (Problems 1 to 22), the *underweighting of rare events* (Problems 23 to 25), and the *loss rate* effect (Problems 26 to 40).

[1] The human data reported in this section were obtained from Ido Erev and Greg Barron.

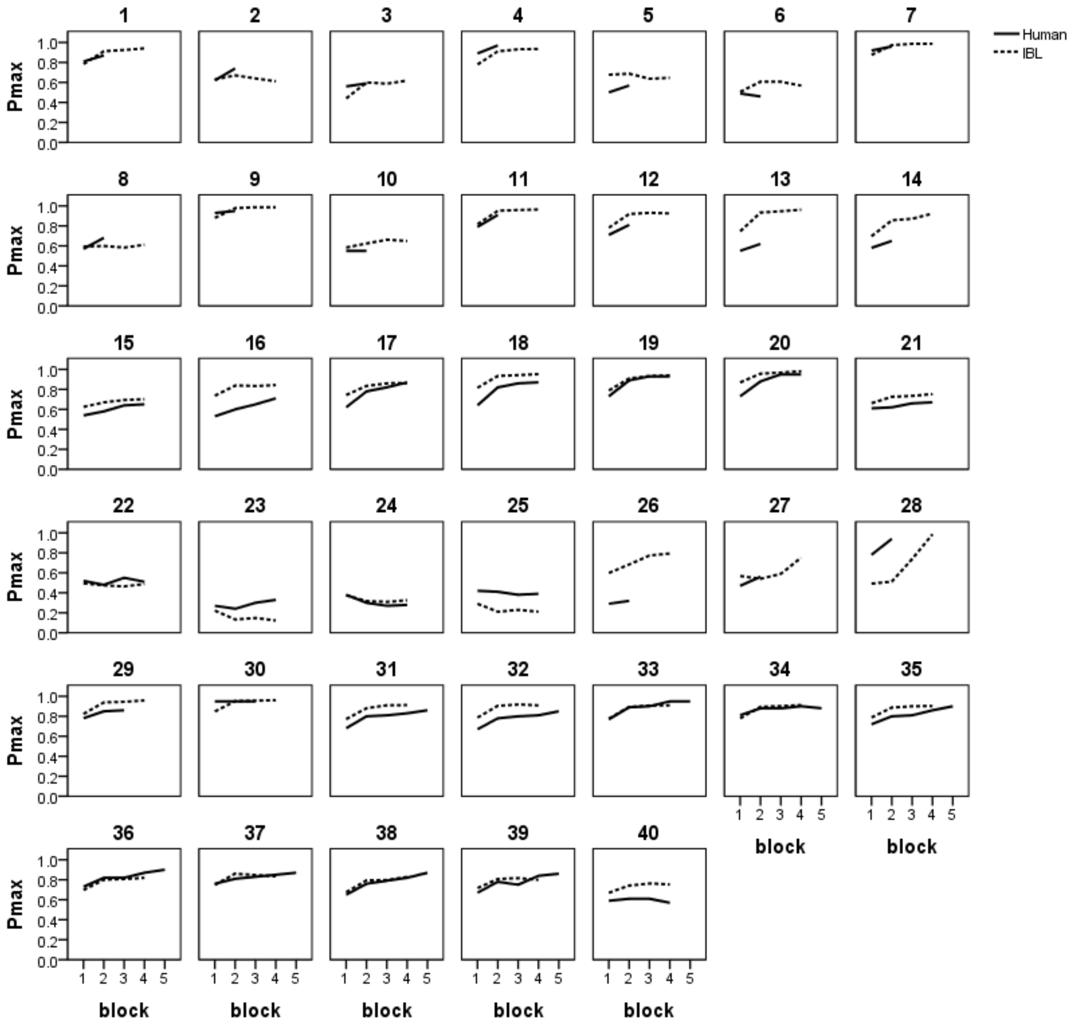

Figure 4. Figure shows learning curves from human data (dark lines) and IBL model data (dotted lines) for each of the 40 problems in Erev and Barron (2005). Each panel represents one of the 40 problems, each problem in the IBL model ran for 400 trials and the panels show the proportion of maximization averaged in blocks of 100 trials. The panels demonstrate the *payoff variability effect* (Problems 1 to 22), the *underweighting of rare events* (Problems 23 to 25), and the *loss rate effect* (Problems 26 to 40).

The source of information for learning in this task is the same as in the generic demonstrations of the TPT data sets described above: the IBL learning mechanisms involving the frequency of observed outcomes, the recency of observed outcomes, and the blended value of the outcomes weighted by the probability of memory retrieval.

Addressing Individual Differences

Erev and Barron (2005) discussed general boundaries of models as predictors, one of them is accounting for individual differences observed in human data. The data generated by the IBL model above is able to capture individual differences found in the problems reported in Erev and Barron (2005). Figure 5 shows the observed distributions of Pmax2 in 32 of the problems (out of the 40 problems shown in Figure 4) for which we had individual data (the black bars). These distributions correspond to the second block (Trials 101-200) over all the participants. Figure 5 also displays the distributions predicted from the IBL model (the white bars). The results show large individual differences in the proportion of maximization in all problems, and remarkably, the same IBL model that predicts the proportion of maximization over time (Figure 4) reproduces the distributions of participants' maximization behavior quite well in the majority of the problems. Although Erev and Barron's RELACS model also produce similar variability in human data, it is worth noting the simplicity of the IBL model compared to RELACS and the generality of the demonstrations from the IBL model compare to those of RELACS.

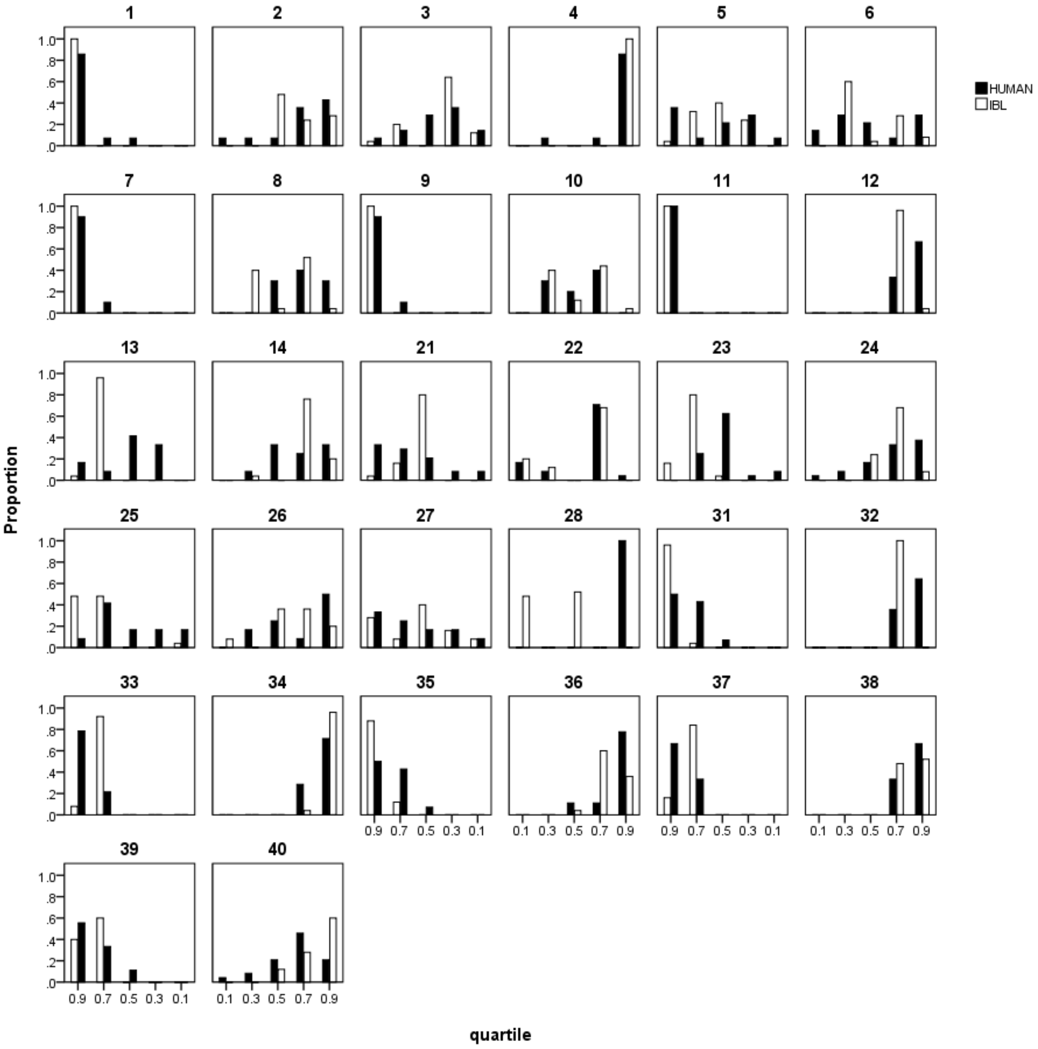

Figure 5. Distribution of proportion of maximization in the second block (Pmax2) over Humans and those produced by the IBL model with the simulated participants in 32 of the 40 problems reported in Erev and Barron (2005) and corresponding to the behavior in Figure 4, for the second block. Each panel represents a problem and the distributions of participants' proportion of maximizations. The y-axis shows the proportion of participants (Humans, dark bars, and simulated by IBL model, white bars).

Probability matching effect

Probability learning refers to the study of how individuals predict the outcome of two mutually exclusive, random events. In a typical probability learning task, participants predict which of two lights will turn on in a number of trials. In the standard version of the task, the probability that a light will turn on is unknown to participants, who learn so from experience. Early studies (Edwards, 1961) suggest a tendency where participants choose the more likely event with a probability that is similar to the event probability, a phenomenon referred to as "probability matching."

Lejarraga et al. (2012) reported the predictions of the IBL model to a set of probability matching problems that were also reported by Erev and Barron (2005) as a test of their RELACS model. The 27 problems were originally taken from Myers et al. (1961). Participants in these experiments had to predict, in each of 150 trials, which of two lights would turn on. Each participant was awarded 100 chips (worth 5¢ each) as game currency, and they could win additional chips by predicting correctly or lose chips by predicting incorrectly. The amount of chips earned at the end of the experiment was exchanged for money. The frequencies of the two lights were 90-10 (i.e., one light turned on 90% of the times and the other light turned on 10% of the times), 70-30, and 50-50. The amount of chips gained with each correct prediction depended on the light being correctly predicted. Because high frequency lights are easier to predict, correct predictions of high frequency lights were rewarded with fewer chips than correct predictions of low frequency lights. There were three gain ratios that determined the rewards: 1:4, 1:2, and 1:1. For example, in the 1:4 condition, correct predictions of low frequency lights were rewarded with 4 chips, while correct predictions of high frequency lights were rewarded with 1 chip. In the 1:1 condition, correct predictions were rewarded with 1 chip irrespective of the lights' frequency. Likewise, because high frequency lights are easier to predict, incorrect predictions for high frequency lights cost more than incorrect predictions for low frequency lights. The cost ratios for incorrect predictions followed the same ratios as for gains. In the 1:4 condition, incorrect predictions of high frequency lights cost 4 chips, while incorrect predictions of low frequency lights cost 1 chip. In the 1:1 condition, incorrect predictions cost 1 chip for both lights. When the two lights occurred with the same frequency (in the 50-50 condition), the light assigned a higher gain was also assigned a lower cost.

The IBL model and the predictions as compared to the results in Myers et al. (1961) were reported in Lejarraga et al. (2012) and reproduced here in Figure 6. The figure shows the mean number of choices for one of the options across participants in each of the 27 problems of Myers et al., (1961). The figure shows accurate predictions of the IBL model (white bars) compared to human data (dark bars) in all the 27 problems.

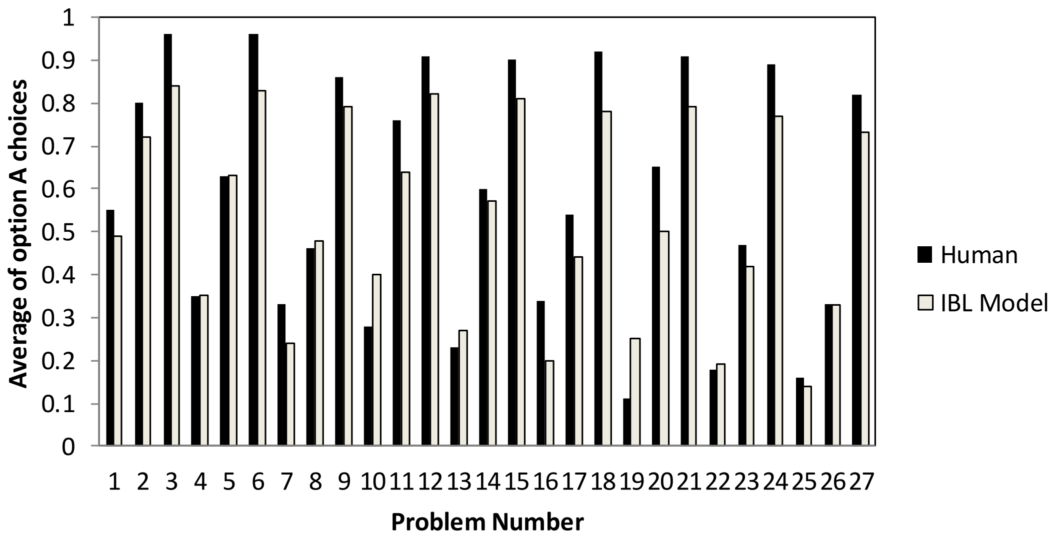

Figure 6. Average choices of option A in 27 problems of Myers et al., (1961) probability learning experiment. The predictions of the IBL model for each problem (white bars) are close to human data (dark bars). For details on the numerical comparison and explanations of the data set see Lejarraga et al. (2012).

Adaptation to nonstationary environments.

Rakow and Miler (2009) explored repeated choice in situations where the outcome probabilities for one of the two options changed over trials. In their

Experiment 1, 40 participants made 100 repeated choices between two risky options in four problems. In all of these problems, each of two options involved a positive and a negative outcome, so participants could win or lose money with each decision. The novelty of the problems studied by Rakow and Miler (2009) is that for one of the options, the probability of the positive outcome remained constant across trials (i.e., the stationary option, S), while this probability changed across trials in the other option (i.e., nonstationary option, NS). Changes in the probabilities for the NS option were gradual: the probability changed .01 per trial and over 40 trials. For example, problem 1 involved a choice between S that offered 10 with a .7 probability or -20 otherwise, and NS that initially offered 10 with a .9 probability or -20 otherwise. From trials 21 to 60, the probability of 10 in NS reduced by .01 in each trial, such that the probability of 10 in trial 60 and onwards was .5. In all four problems, the change in the probability was by .01 per trial and after the 40 changing trials, the probability remained unchanged at .5. After each choice, participants observed the outcome of the chosen option as well as the outcome of the option not chosen (i.e., the foregone payoff). The apparatus and procedures are carefully described in Rakow and Miler (2009). Their results showed that participants adapted slowly to probability changes, a behavior that was not captured particularly well by the associative choice model fitted in that study (Bush & Mosteller, 1955).

We obtained the experimental data from Rakow and Miler (2009) for the four problems in their Experiment 1, and we generated predictions from our IBL model using 100 simulated participants. Detailed results are reported in Lejarraga et al. (2012). Figure 7 shows the IBL model predictions (dotted lines) as compared to the observed data (solid lines), originally reported in Lejarraga et al. (2012).

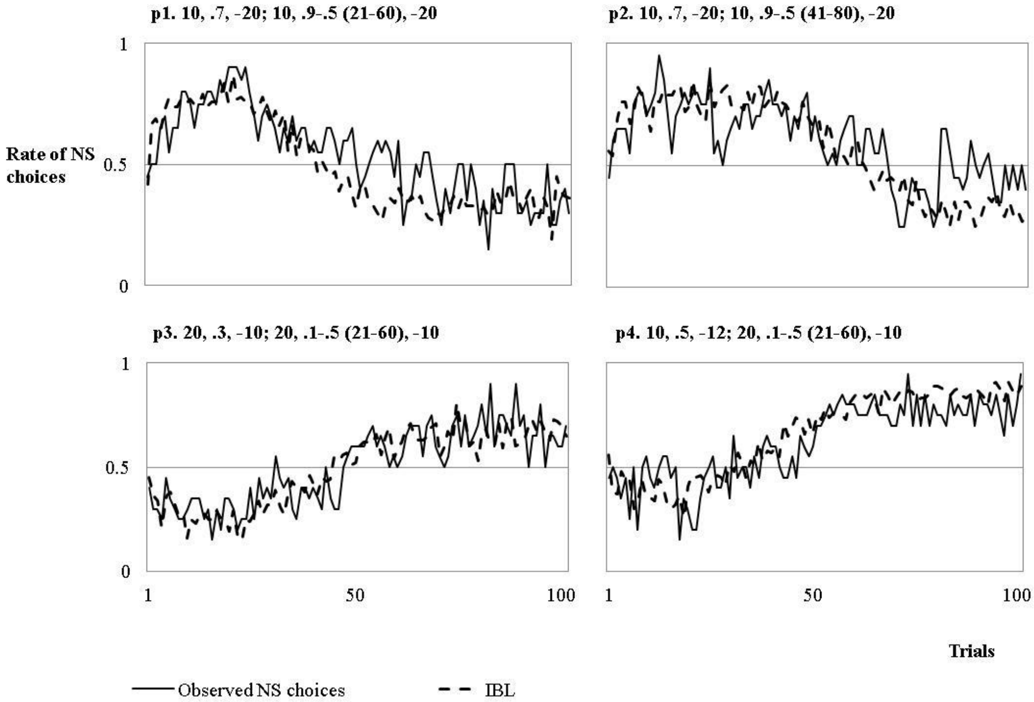

Figure 7. Predictions of the IBL model and human data in four problems designed by Rakow and Miler (2009). Data and tests of the IBL model predictions were reported in Lejarraga et al. (2012).

The accurate predictions of human behavior by the IBL model in all the phenomena demonstrated above support the assertion that the model is an accurate representation of decisions from experience in choice tasks with nonstationary environments. Because the choice problems change gradually across trials, recent experiences are more informative than distant past experiences. In this environment, recency is an adaptive behavior. As Figure 7 shows, participants in Rakow's and Miler's (2009) experiment adapted to changing conditions: Each of the observed learning curves shows a marked change in the trend of choices.

What the IBL model does not explain.

Although the IBL model provides robust predictions across a wide diversity of problems and explains a good number of well-known effects in decisions from experience, the model is not expected to predict behavior accurately in a number of situations. Below there are examples of situations in which the model does not provide accurate predictions. We know there might be many other effects that the model cannot predict and we hope to address the model's miss-predictions in future research.

Pure Risk Aversion

In the demonstrations of the payoff variability effect, Erev and Barron (2005) interpreted the difference between problems 1 and 3 (see Figure 2) as reflecting risk aversion (the high alternative is less attractive when the payoff variability increases), and the difference between problems 1 and 2 as reflecting risk seeking preferences (the low alternative is less attractive when its payoff variability increases). In these problems, however, risk is confounded with expected value, and thus it cannot be interpreted cleanly as a pure risk aversion effect. To explore the pure risk aversion effect, we collected data on a fourth problem not reported in Erev and Barron (2005), in which alternatives are of equal value but they only differ in the variability of the payoff:

Problem 4.	Certain	11	points with certainty
	Risky	21	points with probability 0.5
		1 Otherwise	

Using the same methods as in the first 3 problems, we collected data from 20 participants in problem 4. Results shown in Figure 8 indicate that humans starting at an indifference point (solid line), reduce the proportion of risky choices over time.

The IBL model in contrast (dotted line), starts with a larger preference towards the certain alternative (11) than the risky alternative (21,.5; 1,.5) and moves towards indifference over time. Although the effect is relatively small, the model's trends are in opposition to the humans', and they would be expected to continue in the same direction with even more practice.

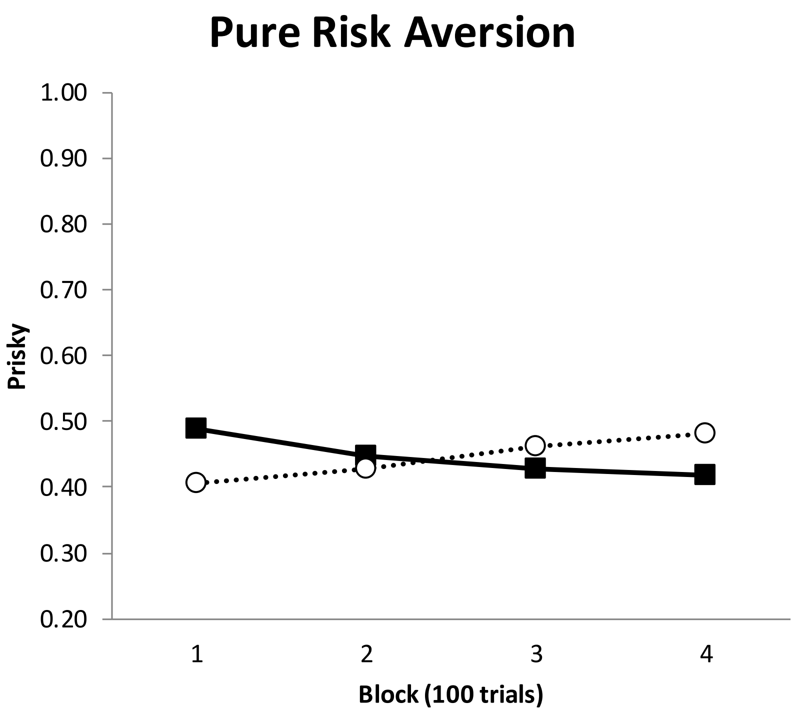

Figure 8. Average human proportions of risky choices (solid line) and the predictions of the IBL model (dotted line), in Problem 4 during 400 trials, averaged in 4 blocks of 100 trials each.

The key insight is that initial experiences of the "1" outcome in the risky option produce a higher blended value for the certain alternative (11) than the risky alternative in the IBL model. The periods in which the risky alternative is selected and the lowest outcome (i.e., 1) is obtained must be longer than the periods of selecting the certain alternative in the first block. Over time, the model "balances out" the value of the two alternatives as experiences of the "21" outcome produce a preference towards the risky alternative.

The question is of course, why do humans and the IBL model differ. The model, building on experiences over time, realizes little by little that the two options have the same expected values through the blended values, and moves towards indifference between the two options. Humans in contrast, seem to maintain and even avoid the "fear" of obtaining a value of "1" that is lower than what they obtain by clicking the safe button "11". This type of "meta-reasoning", beyond reactive decisions based on pure feedback from actions taken are not captured by the IBL model as currently defined. One way in which this initial tendency to "fear" the low outcome of the risky choice might be captured in the model is by creating initial tendencies (higher blended values) for the safe than the risky option.

More risk seeking in losses compared to gain domains

A common effect widely discussed in decisions from description implies that the subjective enjoyment from gaining a certain amount tends to be less than the subjective pain from losing the same amount (Kahneman & Tversky, 1979). Some researchers have demonstrated that *loss aversion* does not hold in decisions from experience, where decision makers seem indifferent between an equal chance of gaining or losing the same amount (Erev, Ert, & Yechiam, 2008; Ert & Erev, 2011). In decisions from description, decision makers are risk averse in the gain domain and risk seeking in the loss domain (Kahneman & Tversky, 1979), and this pattern may reverse or disappear in decisions from experience (Erev & Barron, 2005).

Although much work needs to be done in regards to the differences between gains and losses in decisions from experience, our initial analyses of decisions from experience in the sampling paradigm of the TPT indicate no difference in risky behavior between gains and losses (χ^2= .308, p=.580). The IBL model, however, predicts a difference between gains and losses, which although small, it is significant (χ^2 = 12.462, p<.001). These effects are illustrated in Figure 9. Interestingly, human behavior as well as the IBL model prediction are in disagreement with the predictions from prospect theory: Humans do not show higher risk-seeking tendency in problems involving losses than gains and the IBL model, shows a higher tendency for risky choices in problems involving gains than losses. Both, human data and the IBL model data illustrate opposite effects than those expected in prospect theory.

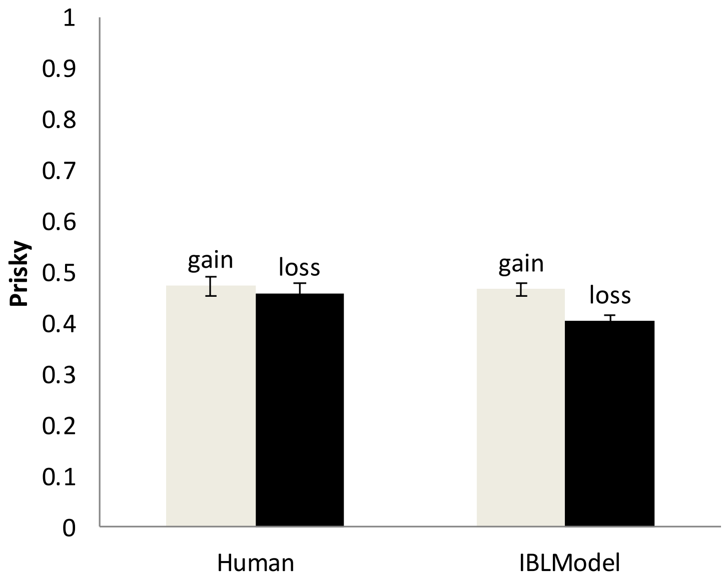

Figure 9. Proportion of risky choices in the gain and loss domains for the TPT sampling paradigm and the predictions of the IBL model.

Emotions, Social, and Non-Cognitive effects

In general, IBLT is a cognitive theory and IBL models are based on memory mechanisms. IBL models are not expected to predict social, emotional, and non-cognitive actions. However, we have started to investigate how the IBL model may account for situations involving two or more individuals involving non-cognitive aspects (e.g., emotions, power, trust). We propose that IBL models may also help in understanding how conflictual social interactions are influenced by the prior experiences of the individuals involved and by the information available to them during the course of interaction (Gonzalez & Martin, 2011). Some initial steps have been taken to use IBL models in multi-person games. For example, Gonzalez and Lebiere (2005) reported a cognitive model for an iterated prisoner's dimella (IPD), initially reported by Lebiere, Wallach, and West (2000), that assumes instances are stored in memory, including one's own action, the other player's action, and the payoff. More recently, the IBL model was used in more complex multi-person task, the market entry game (Gonzalez et al., 2011). This model, which obtained the runner-up prize in a modeling competition, shares basic features with IBL models of individual choice (e.g., Lejarraga et al., 2012), and importantly no explicit modifications were included in the model to account for the effects of the market entry task.

Many models of individual decisions from experience are incapable of representing human behavior in social contexts. For example, Erev and Roth (2001) noted that simple reinforcement learning models predicted the effect of experience in two-person games like the Iterated Prisoner's Dilemma (IPD) only in situations where players could not punish or reciprocate. A simple model predicts a decrease in cooperation over time, even though most behavioral experiments demonstrate an increase in mutual cooperation due to the possibility of reciprocation (Rapoport & Chammah, 1965; Rapoport & Mowshowitz, 1966). To account for the effects of reciprocation, Erev and Roth (2001) made two explicit modifications to the basic reinforcement learning model: if a player adopts a reciprocation strategy, he will cooperate in the next trial only if the other player has cooperated in the current trial; the probability that a player continues to do so will depend on the number of times the reciprocation strategy was played. Although these tweaks to the model may accurately represent the kind of cognitive reasoning that people actually use in the IPD, they are unlikely to generalize to other situations with different action sets or outcomes. The IBL model appears to account for these reciprocity effects without the need for explicit and situation-specific rules (Gonzalez, Dutt, Martin, & Ben-Asher, 2012; in preparation; Juvina et al., 2011). However, much work is needed for understanding how the IBL model can be extended to account for the effect of non-cognitive variables (e.g., emotions, social considerations such as power, fairness, envy, etc.) on decision making.

Conclusions

Research on decisions from experience has demonstrated great potential to expand our understanding of the processes involved in making decisions. Experimental and cognitive modeling approaches to study of experience-based choice help open a window to understanding processes beyond the observable choice. With simple experimental paradigms, researchers have improved our understanding of the processes that lead to a choice, such as the recognition of alternatives, the formation of preferences, the evaluation of outcomes, the integration of experiences and the projection of costs and benefits. With cognitive models, researchers have helped to explain how these processes develop, and to predict behavior in some novel circumstances.

A problem, which I have aimed to address in the past years, is the lack of a comprehensive model for experience-based choice behavior and the proliferation of task-specific models of decisions from experience. Several on-going efforts have addressed this issue in many different ways through comprehensive model comparison and demonstrations (Gonzalez & Dutt, 2011; Lejarraga et al., 2012), and through model prediction competitions (Erev, Ert, & Roth, 2010; Erev et al., 2010). These efforts are converging over how decisions from experience are explained: via cognitive memory processes, including recency and frequency of events. Our explanations come from models based on IBLT that have shown robust and accurate predictions in multiple tasks.

This chapter summarizes the history of IBLT and IBL models. Furthermore, it highlights and attempts to start addressing an important problem in this research program: the robustness and specificity tradeoff. Although the IBL models have shown robustness and generality, they also need to clearly and more specifically guide the potential users of these models to explain concrete phenomena in decision sciences. We summarized some phenomena that the IBL model explains: payoff variability effect, underweighting of rare events, loss rate effect, individual differences, probability matching, and adaptation to nonstationary environments. We also summarized some phenomena that the model in its current form is unable to capture: the pure risk aversion effect, more risk seeking in losses compared to in gains domains, and emotions, social, and non-cognitive effects. Future research will address these and many other challenges that the IBL model faces.

References

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

Brehmer, B. (1990). Strategies in real-time, dynamic decision making. In R. M. Hogarth (Ed.), *Insights in decision making* (pp. 262-279). Chicago: University of Chicago Press.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica, 81*(3), 211-241.

Busemeyer, J. R., & Wang, Y. M. (2000). Model comparison and model selections based on generalization criterion methodology. *Journal of Mathematical Psychology, 44*(1), 171-189.

Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. Oxford, UK: John Wiley & Sons, Inc.

Dienes, Z., & Fahey, R. (1995). Role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(4), 848-862.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: The Free Press.

Dutt, V., Yamaguchi, M., Gonzalez, C., & Proctor, R. W. (2009). An instance-based learning model of stimulus-response compatibility effects in mixed location-relevant and location-irrelevant tasks. In A. Howes, D. Peebles & R. Cooper (Eds.), *9th International Conference on Cognitive Modeling – ICCM2009*. Manchester, UK.

Edwards, W. (1961). Behavioral decision theory. *Annual Review of Psychology, 12,* 473-498.

Edwards, W. (1962). Dynamic decision theory and probabilistic information processing. *Human Factors, 4*, 59-73.

Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review, 112*(4), 912-931.

Erev, I., Ert, E., & Roth, A. E. (2010). A choice prediction competition for market entry games: An introduction. *Games, 1*, 117-136.

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S., Hau, R., et al. (2010). A choice prediction competition for choices from experience and from description. *Journal of Behavioral Decision Making, 23*, 15-47.

Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making, 21*, 575-597.

Erev, I., & Roth, A. E. (2001). Simple reinforcement learning models and reciprocation in the Prisoner's Dilemma game. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (pp. 215-231). Cambridge, MA: MIT Press.

Ert, E., & Erev, I. (2011). On the descriptive value of loss aversion in decisions under risk. Unpublished manuscript.

Gibson, F. P., Fichman, M., & Plaut, D. C. (1997). Learning in dynamic decision tasks: Computational model and empirical evidence. *Organizational Behavior and Human Decision Processes, 71*(1), 1-35.

Gilboa, I., & Schmeidler, D. (1995). Case-based decision theory. *The Quarterly Journal of Economics, 110*(3), 605-639.

Gilboa, I., & Schmeidler, D. (2000). Case-based knowledge and induction. *IEEE Transactions on Systems, Man, and Cybernetics-PART A: Systems and Humans, 30*(2), 85-95.

Gonzalez, C. (2004). Learning to make decisions in dynamic environments: Effects of time constraints and cognitive abilities. *Human Factors, 46*(3), 449-460.

Gonzalez, C. (2005). Decision support for real-time dynamic decision making tasks. *Organizational Behavior & Human Decision Processes, 96*(2), 142-154.

Gonzalez, C., Best, B. J., Healy, A. F., Bourne, L. E., Jr, & Kole, J. A. (2010). A cognitive modeling account of simultaneous learning and fatigue effects. *Journal of Cognitive Systems Research, 12*(1), 19-32.

Gonzalez, C., & Dutt, V. (2010). Instance-based learning models of training. In *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 2319-2323). San Francisco, CA: Human Factors and Ergonomics Society.

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating decisions from experience in sampling and repeated choice paradigms. *Psychological Review, 118*(4), 523-551.

Gonzalez, C., Dutt, V., & Lebiere, C. (in press). Validating instance-based learning mechanisms outside of ACT-R. *Journal of Computational Science*.

Gonzalez, C., Dutt, V., & Lejarraga, T. (2011). A loser can be a winner: Comparison of two instance-based learning models in a market entry competition. *Games, 2*(1), 136-162.

Gonzalez, C., Dutt, V., Martin, J., & Ben-Asher N. (2012). Decisions from Experience in Conflict Situations: Cognitive Model of the Effects of Interdependence Information. *Behavioral Decision Research in Management*

Conference. BDRM 2012. June 27-29, 2012. Leeds School of Business, Boulder, CO

Gonzalez, C., & Lebiere, C. (2005). Instance-based cognitive models of decision making. In D. Zizzo & A. Courakis (Eds.), *Transfer of knowledge in economic decision-making* (pp. 148-165). New York: Macmillan (Palgrave Macmillan).

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

Gonzalez, C., & Martin, J. M. (2011). Scaling up Instance-Based Learning Theory to account for social interactions. *Negotiation and Conflict Management Research, 4*(2), 110-128.

Haruvy, E., & Erev, I. (2001). On the application and interpretation of learning models. In R. Zwick & A. Rapoport (Eds.), *Advances in experimental business research* (pp. 285-300). Norwell, MA: Kluwer Academic.

Hertwig, R. (2012). Tapping into the wisdom of the crowd - with confidence. *Science, 336*(303-304).

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15*(8), 534-539.

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences, 13*(12), 517-523.

Juvina, I., Lebiere, C., Martin, J. M., & Gonzalez, C. (2011). Intergroup prisoner's dilemma with intragroup power dynamics. *Games, 2*, 21-51.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-291.

Kerstholt, J. H., & Raaijmakers, J. G. W. (1997). Decision making in dynamic task environments. In R. Ranyard, W. R. Crozier & O. Svenson (Eds.), *Decision making: Cognitive models and explanations* (pp. 205-217). London: Routledge.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.). (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex Publishing Corporation.

Lebiere, C. (1998). *The dynamics of cognition: An ACT-R model of cognitive arithmetic*. Unpublished Ph.D. Dissertation, Carnegie Mellon University, Pittsburgh, PA.

Lebiere, C., Gonzalez, C., & Martin, M. (2007). Instance-based decision making model of repeated binary choice. In R. L. Lewis, T. A. Polk & J. E. Laird (Eds.), *Proceedings of the 8th International Conference on Cognitive Modeling* (pp. 67-72). Oxford, UK: Psychology Press.

Lebiere, C., Wallach, D., & West, R. L. (2000). A memory-based account of the prisoner's dilemma and other 2x2 games *Proceedings of the International Conference on Cognitive Modeling* (pp. 185-193). NL: Universal Press.

Lejarraga, T., Dutt, V., & Gonzalez, C. (2010). *Instance-based learning in repeated binary choice*. Paper presented at the Society for Judgement and Decision Making.

Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making, 25*(2), 143-153.

Martin, M. K., Gonzalez, C., & Lebiere, C. (2004). Learning to make decisions in dynamic environments: ACT-R plays the beer game. In M. C. Lovett, C. D.

Schunn, C. Lebiere & P. Munro (Eds.), *Proceedings of the Sixth International Conference on Cognitive Modeling* (Vol. 420, pp. 178-183). Pittsburgh, PA: Lawrence Erlbaum Associates Publishers.

Myers, J. L., Reilly, R. E., & Taub, H. A. (1961). Differential cost, gain, and relative frequency of reward in a sequential choice situation. *Journal of experimental Psychology, 62*(4), 357-360.

Pew, R. W., & Mavor, A. S. (1998). *Modeling human and organizational behavior*. Washington: National City Press.

Rakow, T., & Miler, K. (2009). Doomed to repeat the successes of the past: History is best forgotten for repeated choices with nonstationary payoffs. *Memory & Cognition, 37*(7), 985-1000.

Rapoport, A., & Chammah, A. M. (1965). *Prisoner's dilemma: A study in conflict and cooperation*. Ann Arbor: University of Michigan Press.

Rapoport, A., & Mowshowitz, A. (1966). Experimental studies of stochastic models for the Prisoner's dilemma. *System Research and Behavioral Science, 11*(6), 444-458.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131.

Zsambok, C. E., & Klein, G. A. (Eds.). (1997). *Naturalistic decision making*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.