Psychological Review

Instance-Based Learning: Integrating Sampling and Repeated Decisions From Experience

Cleotilde Gonzalez and Varun Dutt

Online First Publication, August 1, 2011. doi: 10.1037/a0024558

Psychological Review

0033-295X/11/$12.00 DOI: 10.1037/a0024558

Instance-Based Learning: Integrating Sampling and Repeated Decisions From Experience

Cleotilde Gonzalez and Varun Dutt
Carnegie Mellon University

In decisions from experience, there are 2 experimental paradigms: sampling and repeated-choice. In the sampling paradigm, participants sample between 2 options as many times as they want (i.e., the stopping point is variable), observe the outcome with no real consequences each time, and finally select 1 of the 2 options that cause them to earn or lose money. In the repeated-choice paradigm, participants select 1 of the 2 options for a fixed number of times and receive immediate outcome feedback that affects their earnings. These 2 experimental paradigms have been studied independently, and different cognitive processes have often been assumed to take place in each, as represented in widely diverse computational models. We demonstrate that behavior in these 2 paradigms relies upon common cognitive processes proposed by the instance-based learning theory (IBLT; Gonzalez, Lerch, & Lebiere, 2003) and that the stopping point is the only difference between the 2 paradigms. A single cognitive model based on IBLT (with an added stopping point rule in the sampling paradigm) captures human choices and predicts the sequence of choice selections across both paradigms. We integrate the paradigms through quantitative model comparison, where IBLT outperforms the best models created for each paradigm separately. We discuss the implications for the psychology of decision making.

Keywords: instance-based learning, decisions from experience, sampling, repeated-choice, quantitative model comparison

In many real-world situations, we have an opportunity to sample options before making decisions for real: Dressing rooms allow people to try on different garments before purchasing; ice cream stores often give samples; and wine tasting is a good way to evaluate wine, especially when bottles sell for high prices. Similarly, it is wise to see a good number of houses before deciding to purchase one. In fact, a recent TV show made it possible for a buyer to "sleep on it," live in the house for 24 hr, and "experience" the good and bad of the house before making a purchase (HGTV show, *Sleep on It*).

Unfortunately, we do not always have such luxuries of sampling. In many situations, decision makers may need to "take it or leave it" and be forced to make consequential decisions without exploring the options. While driving in a new city for example, approaching intersections involves choices that need to be made in real-time and the consequences observed afterwards. Traders in the stock market need to be able to make sense of a large amount of data, interpreting rapid fluctuations of stocks and reacting without delays, which could cost thousands of dollars. Firefighting commanders often need to make difficult decisions in the face of rapidly growing fire. In all these situations, decision makers would not have a chance to sample before making a consequential choice. The *Sleep on It* TV show might be good fun, but it is not practical and is not a common practice in the real-estate business. That is, many times one needs to make consequential decisions and learn from the outcomes.

Although different, the two types of real-world decision-making situations described above have something in common: They all require decision makers to act based on their own experience (*decisions from experience*; DFE). These real-world situations represent two main types of decision making paradigms used to study DFE (Hertwig & Erev, 2009): sampling and repeated-choice. These two paradigms have been studied independently, and different cognitive processes are often assumed to take place in each. The main goal of this research is to demonstrate that the two paradigms have common cognitive processes explained by instance-based learning theory (IBLT; Gonzalez, Lerch, & Lebiere, 2003), a theory of DFE in dynamic tasks.

Common Behavioral Patterns of DFE in Contrast to Decisions From Description

DFE represents a major recent breakthrough in behavioral decision research and a shift of attention away from the study of "decisions from description," where choices are made from explicitly stated payoffs and associated probabilities (Barron & Erev,

Cleotilde Gonzalez and Varun Dutt, Dynamic Decision Making Laboratory, Department of Social and Decision Sciences, Carnegie Mellon University.

This research was supported in part by a grant from the National Science Foundation (Human and Social Dynamics: Decision, Risk, and Uncertainty, Award 0624228) and by a grant from the Defense Threat Reduction Agency (HDTRA1-09-1-0053) to Cleotilde Gonzalez. We thank Hau-yu Wong, Tomás Lejarraga, and Jolie Martin for their review of this article; Ido Erev and Ralph Hertwig for making the human data available that was essential for conducting this research; and the authors of the models reported in this article who provided the code for our analyses.

Correspondence concerning this article should be addressed to Cleotilde Gonzalez, Dynamic Decision Making Laboratory, Carnegie Mellon University, Pittsburgh, PA 15213. E-mail: coty@cmu.edu

2003; Hertwig, Barron, Weber, & Erev, 2004). Prospect theory (Kahneman & Tversky, 1979) has been a prominent model to explain the deviations from expected utility theory in decisions from description. Most research on DFE to this date has focused on finding and explaining a behavioral "gap" between decisions made from description and experience called the *description–experience gap* (Barron & Erev, 2003; Hertwig et al., 2004; Hertwig & Erev, 2009). Researchers have demonstrated that in the presence of rare outcomes, decisions from description and DFE lead to dramatically different choice behavior. In description, people behave as if rare outcomes receive more impact than they deserve according to their objective probability, whereas in experience, people behave as if rare outcomes receive less impact than they deserve (Hertwig & Erev, 2009). For example, consider a situation where participants select between two options, one high and the other low:

(a) \$32 ($p = 0.1$) or \$0 ($1 - p = 0.9$)

(b) \$3 ($p = 1$)

The high option (a) has a higher expected value (\$3.2) than the low option (b) (\$3), but the high outcome (\$32) occurs with low probability (0.1). In this situation, when making decisions from the description, participants behave as if the low probability outcome (\$32) has more impact than it deserves (the proportion of maximization is about 48%; Hertwig et al., 2004). In contrast, participants behave as if the low probability outcome (\$32) has less impact than it deserves in DFE (the proportion of maximization is only about 20% in the sampling paradigm and about 24% in the repeated-choice paradigm; Hertwig et al., 2004).

The behavioral pattern described above is consistent in both DFE paradigms. Behavior in sampling and repeated-choice paradigms is highly correlated ($r = .93$, $p < .01$; Hertwig et al., 2004), and there is a consistent description–experience gap in the sampling paradigm (Hertwig et al., 2004) and in the repeated-choice paradigm (Barron & Erev, 2003). Even more broadly, the data collected for the *Technion Prediction Tournament* (TPT), a modeling competition that focused on three experimental paradigms (decisions from description, decisions by sampling, and repeated-choice) and two large data sets (estimation and competition), demonstrated the robustness of the description–experience gap and the similarity of behavior in the DFE paradigms (Erev, Ert, Roth, et al., 2010). A correlation analysis of the estimation set reported in Erev, Ert, Roth, et al. (2010) indicated similar negative correlations between choices made in the description and sampling paradigms ($r = -.53$, $p = .0004$) and between choices made in the description and repeated-choice paradigms ($r = -.37$, $p = .004$), whereas the correlation between choices made in the sampling and repeated-choice paradigms was positive and highly significant ($r = .83$, $p < .0001$). Regardless of the behavioral similarities in sampling and repeated-choice, very different models won the competition in the DFE paradigms (Erev, Ert, Roth, et al., 2010), and, thus, widely different cognitive assumptions have prevailed across the two paradigms.

Integrating DFE Paradigms Through Quantitative Model Comparison

In contrast to decisions from description where prospect theory (Kahneman & Tversky, 1979) has been a prominent model to explain human-choice behavior, a model that can explain DFE across paradigms and tasks has not yet been found. In fact, a challenge in understanding the cognitive processes involved in DFE is the proliferation of highly task-specific cognitive models that often predict behavior in a particular task but fail to explain behavior even in other closely related tasks (Lejarraga, Dutt, & Gonzalez, 2010). Our main argument in this article is that common behavioral patterns in the sampling and repeated-choice paradigms are well represented, explained, and generalized by a single computational model based on IBLT (hereafter, IBL model; Gonzalez et al., 2003). We demonstrate how one IBL model captures human behavior in both DFE paradigms better than individual models that have been created to account for choices in each paradigm separately.

To integrate the DFE paradigms, we follow a quantitative model comparison approach. We compare the more prominent models that have been created exclusively under each paradigm to the IBL model on how well they account and predict human behavior. Because there are several diverse models with distinctive complexities within each paradigm, we follow goodness of fit and generalizability procedures to compare them (Pitt & Myung, 2002; Pitt, Myung, & Zhang, 2002). We also use available human data sets that share the same choice problems across both paradigms. We bring together *six choice* (SC) problems used in two separate studies, the repeated-choice paradigm (Barron & Erev, 2003) and the sampling paradigm (Hertwig et al., 2004); and we use the sampling and repeated-choice data sets from the TPT (Erev, Ert, Roth, et al., 2010). Following calibration and generalizations procedures, we show that the IBL model predicts general choice behavior across both paradigms—better than the winning models in the TPT and other models developed exclusively to account for behavior in each paradigm separately.

Furthermore, we go beyond predicting a general proportion of risky choices (R-rate). We demonstrate that, in addition to predicting the R-rate, the IBL model explains and predicts the sequence of selections of the two options (i.e., the rate of alternations from one option to the other, or the A-rate) made in the sampling and repeated-choice paradigms. Thus, the A-rate is predicted by the same cognitive mechanisms proposed by IBLT. The only additional mechanism needed in the IBL model to account for the sampling sequence is a variable *stopping point* rule, which demarks when the sampling process ends before a final decision is made and is fixed in the repeated-choice paradigm.

Together with existent evidence of the wide-ranging applications of IBLT to DFE in different contexts (Gonzalez & Dutt, 2010; Gonzalez & Lebiere, 2005; Gonzalez et al., 2003; Lejarraga et al., 2010), the current research demonstrates the generality of IBLT to explain and predict DFE in the DFE paradigms. Moreover, the general learning mechanisms of IBLT are a variant of some of the proposed mechanisms in the Adaptive Control of Thought—Rational (ACT–R) theory of cognition (Anderson & Lebiere, 1998), which were originally developed to capture decisions in dynamic settings (Gonzalez et al., 2003).

In what follows, we first present a discussion of the cognitive explanations for the common effects and the differences found in the sampling and repeated-choice paradigms. Next, we summarize IBLT and present a concrete IBL model used in this research. We then summarize existent models that account for behavior distinctly in each paradigm. Following this section, the IBL model is compared to the best models known to account for human behavior in each paradigm by using a set of calibration and generalization procedures. Finally, we discuss the psychological interpretations of our findings, the implications and limitations of the IBL model, and the broader implications of this work to the psychology of decision making.

Common Behavioral Patterns of Sampling and Repeated-Choice Paradigms Beyond the Description–Experience Gap

Behavioral decision researchers have used a simple experimental tool where respondents are presented with two buttons on a computer screen (Barron & Erev, 2003; Hertwig & Erev, 2009). Each button represents a payoff distribution unknown to participants. Clicking a button results in a random draw from the distribution, and an outcome is presented. In the sampling paradigm (Hertwig et al., 2004), participants are asked to choose between the two buttons as many times as they want (i.e., the stopping point is variable) and observe the outcomes with no real consequences, before they make a single choice between the two options that causes them to actually earn or lose money. In the repeated-choice paradigm (Barron & Erev, 2003), participants are asked to make a fixed number of choices between the two options (i.e., the stopping point is fixed); each choice affects their earnings, and they receive immediate feedback on obtained outcomes (the total number of choices is not told to participants).

As explained above, there are common behavioral patterns found in the DFE paradigms in contrast to decisions from description, and only recently have possible differences been discussed (Camilleri & Newell, 2011; Hertwig & Erev, 2009; Rakow & Newell, 2010). In these reviews, the cognitive explanations for the DFE paradigms' common behavioral effects are often different, and in some cases the paradigms are considered to be different without explaining the common behavioral effects (Camilleri & Newell, 2011). In the sampling paradigm, the leading cognitive explanation of the description–experience gap has been "reliance on small samples," and "reliance on recent experiences of outcomes" in the repeated-choice paradigm. Across a number of studies, it was found that people sample relatively few times (median number of samples often vary between 11 and 19; Hau, Pleskac, Kiefer, & Hertwig, 2008; Hertwig & Erev, 2009), reducing the chances of experiencing the rare outcome. Because there is no sampling and each choice affects a participant's earnings in repeated-choice, the favored explanation of reliance on recent experiences of outcomes implies giving more weight to more recent events. According to this recency argument, rare outcomes are underweighted because people remember more recent events better than less recent events, and rare outcomes are less likely to have occurred recently (Erev & Barron, 2005; Fiedler, 2000; Kareev, 2000).

Another, perhaps less well-understood, cognitive explanation that is claimed to be different between the sampling and repeated-choice paradigms is the exploration–exploitation tradeoff: Selecting new options may provide the possibility of greater rewards and learning at the risk of low outcomes (exploration), whereas repeating previous choices is likely to yield known outcomes at the risk of lacking the opportunity for improvement (exploitation). Many researchers assume that the exploration–exploitation tradeoff does not exist in the sampling paradigm because these two stages occur separately: Exploration occurs during sampling, whereas exploitation occurs when the choice is made for real (Camilleri & Newell, 2011; Rakow & Newell, 2010). Further, researchers often assume a random choice and high exploration rate during sampling (Rakow & Newell, 2010).

In contrast, researchers assume that the exploration-exploitation tradeoff exists in repeated-choice because these two phases are difficult to disentangle. Because the choices are consequential, people exhibit a "hot-stove" effect, where high outcomes increase the probability of repeating a choice (exploitation), and low outcomes decrease the probability (exploration; Biele, Erev, & Ert, 2009; Denrell & March, 2001). Recent empirical evidence suggests that consequential choices in the repeated-choice paradigm may produce stronger underweighting of low probability outcomes and a tendency to be biased away from risky options due to the hot-stove effect, whereas this effect may be absent in the sampling paradigm because sampling does not have a cost (Camilleri & Newell, 2011). More recent studies have demonstrated that positively surprising obtained payoffs and negatively surprising forgone payoffs reduce *inertia* (the rate of repeating the previous choice) in the repeated-choice paradigm (Nevo & Erev, 2011).

Similar Exploration–Exploitation Transition Between the Two DFE Paradigms

Because the similarities between the DFE paradigms have only been investigated through the general proportion of risky choices (R-rate), the commonality in exploration and exploitation between the paradigms has been overlooked up to now. A measure of exploration in the DFE paradigms is the proportion of alternations (A-rate): switches from one option to the other (Erev, Ert, & Roth, 2010). When we analyzed the A-rate in the human data collected in the TPT (Erev, Ert, Roth, et al., 2010), we observed that overall the average A-rate in the sampling paradigm is nearly double (0.34 in the estimation set and 0.29 in the competition set) of that in the repeated-choice paradigm (0.14 in the estimation set and 0.13 in the competition set). The proportion of risky choices (R-rate), however, is about the same in both paradigms (0.49 in sampling and 0.40 in repeated-choice for the estimation set; and 0.44 in sampling and 0.38 in repeated-choice for the competition set). Thus, although the paradigms are similar at general choice level (in the R-rate), they appear to be very different at the average process (A-rate) level. A possible explanation for the higher A-rate in the sampling paradigm is that sampling does not cost, whereas choices in repeated-choice are costly. However, it has been recently found that the A-rate varies widely in the sampling paradigm, where some participants are frequent and others are infrequent alternators (Hills & Hertwig, 2010). Thus, the fact that sampling is costless does not explain the diversity of sampling strategies and the reliance on small samples.

If the same cognitive mechanisms of IBLT are to explain not only the R-rate but also the A-rate across both paradigms, we expect a high correlation of the A-rate between sampling and repeated-choice paradigms across trials (samples). Furthermore, given the reliance on small samples in the sampling paradigm, we expect this correlation to be higher for smaller sample sizes and lower for larger sample sizes (because fewer participants engage in large samples). We calculated these correlations in the TPT data sets (Erev, Ert, Roth, et al., 2010). The correlations of the A-rate between the sampling and repeated-choice paradigms for the first nine samples (median) are high and highly significant: $r = .93$, $p < .01$, for the estimation set, and $r = .89$, $p < .01$, for the competition set. Furthermore, as expected, these correlations decrease as the number of trials used in the calculation increase. Thus, for the first 99 samples, $r = .35$, $p < .05$, for the estimation set, and $r = .70$, $p < .01$, for the competition set. These results demonstrate the highly similar sequence of alternations between the two options in both paradigms and a source of differences between the two paradigms: the stopping point in the sampling paradigm.

IBLT

IBLT was developed to explain and predict learning and decision making in real-time dynamic tasks (Gonzalez et al., 2003). Dynamic tasks are characterized by decision conditions that change spontaneously and as a result of previous decisions while attempting to maximize gains over the long run (Edwards, 1962; Rapoport, 1975). Dynamic tasks range in their levels of dynamic characteristics (Edwards, 1962). The least dynamic tasks involve sequential decisions in an environment where neither the environment nor the participants' information about it is affected by their previous decisions. The most dynamic tasks involve sequential decisions in situations where the environment and the participants' information about it changes over time and as a function of participants' previous decisions (Edwards, 1962). IBLT represents the process by which decisions are made through experience in a range of dynamic tasks, from the least dynamic to the most dynamic ones. Both the sampling and repeated-choice paradigms are examples of less dynamic tasks.

IBLT proposes a key representation of cognitive information: an instance. An instance is a representation of each decision option, often consisting of three parts: a situation (a set of attributes that define the option), a decision for one of the many options, and an outcome resulting from making that decision. The theory also proposes a generic decision-making process that starts by recognizing decision situations, generating instances through the interaction with the decision task, and finishes with the reinforcement of instances that led to good decision outcomes through feedback. The general decision-making process is explained in detail in Gonzalez et al. (2003), and it involves the following steps: the recognition of a situation from a task and the creation of decision options; the retrieval of instances from memory that are similar to the current task's situation, or the use of decision heuristics in the absence of similar instances in memory; the selection of the best decision option; and the process of reinforcing instances corresponding to observed outcomes through a process of feedback.

IBL models are particular representations of IBLT for specific tasks. Many IBL models have been developed in a wide variety of tasks, including dynamically complex tasks (Gonzalez & Lebiere, 2005; Martin, Gonzalez, & Lebiere, 2004), training paradigms of simple and complex tasks (Gonzalez, Best, Healy, Kole, & Bourne, 2011; Gonzalez & Dutt, 2010), simple stimulus–response practice and skill acquisition tasks (Dutt, Yamaguchi, Gonzalez, & Proctor, 2009, 2010), and repeated binary-choice tasks (Lebiere, Gonzalez, & Martin, 2007; Lejarraga et al., 2010), among others. The different applications of IBLT illustrate its generality and its ability to explain learning from exploration and DFE in multiple tasks and contexts. Most of the current IBL models, however, are task-specific.

A recent IBL model has shown that generalization of the theory is possible across multiple tasks that share the same task structure. The model reported in Lejarraga et al. (2010) was built to predict performance in individual repeated binary-choice tasks, probability-learning tasks, and repeated-choice tasks with changing probability of outcomes as a function of trials. In this article, we use an extension of the IBL model reported in Lejarraga et al. to explain and predict performance across the two DFE paradigms.

The IBL Model of Sampling and Repeated-Choice

Instances in the DFE's two-button task have a much simpler representation compared to other IBL models. The instance structure is simple, because the task structure in the two paradigms is also simple. Each instance consists of a label that identifies a decision option in the task and the outcome obtained. For example, (Right, \$4) is an instance where the decision was to click the button on the right side, and the outcome obtained was \$4. The instances in the sampling and repeated-choice paradigms are treated exactly in the same way.

The model we report here builds on the one reported in Lejarraga et al. (2010) by adding an inertia rule (Equation 1 below; built on the work of Biele et al., 2009). This rule determines whether the previous choice in the task is repeated according to a random draw from a uniform distribution. If the previous choice is not repeated, then the option with the highest utility (*blended* value; Equation 2 below) is selected. An option's blended value depends on its associated outcomes and the probability of retrieving corresponding instances from memory (Equation 3 below). Furthermore, the probability of retrieving instances from memory is a function of their *activation* (Equation 4 below). Activation in the model is a function of the recency and frequency of retrieving instances from memory (a simplification of the full activation mechanism developed in the ACT–R cognitive architecture; Anderson & Lebiere, 1998, 2003).

Inertia mechanism. We build on the work of Biele et al. (2009) and use a free parameter in the IBL model, the probability of inertia (*pInertia*), to determine whether the choice made in the previous trial is repeated or not. A choice is made in the model in trial $t + 1$ as follows:

> *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). (1)

pInertia could vary between 0 and 1, and it does not change across trials or participants. Naturally, the higher the value of *pInertia*, the more frequently the IBL model will repeat the last choice in repeated trials or samples.

Blending and activation mechanisms. If the most recent decision is not repeated according to Equation 1, then the model selects an option with the highest blended value V (Gonzalez et al., 2003; Lejarraga et al., 2010) from all instances with outcomes that belong to that option. The blended value of option j is defined as

$$V_j = \sum_{i=1}^{n} p_i x_i \tag{2}$$

Where x_i is the value of the observed outcome in the outcome slot of an instance i corresponding to the option j, and p_i is the probability of that instance's retrieval from memory (for the case of a binary-choice task, the value of j in Equation 2 could be either safe or risky, or maximizing or minimizing). The blended value of an option is the sum of all observed outcomes x_i in the outcome slot of corresponding instances, weighted by the instances' probability of retrieval. Thus, the blended value in this model is a form of cognitive representation of experienced utility, given the use of only observed outcomes and the reliance on human memory mechanisms.

In any trial t, the probability of retrieving instance i from memory is a function of that instance's activation relative to the activation of all other instances corresponding to that option, given by

$$P_{i,t} = \frac{e^{A_{i,t}/\tau}}{\sum_j e^{A_{j,t}/\tau}} \tag{3}$$

Where τ is random noise defined as $\sigma \times \sqrt{2}$, and σ is a free noise parameter. Noise parameter σ captures the imprecision of retrieving instances from memory. The noise parameter has no default value in the ACT–R architecture from which it was borrowed (Anderson & Lebiere, 1998); however, the parameter has been found to have a mean of 0.45 in many ACT–R studies (Wong, Cokely, & Schooler, 2010). Higher σ values imply a greater variability in the retrieving instances from memory.

The activation of each instance in memory depends upon the *activation* mechanism originally proposed in ACT–R (Anderson & Lebiere, 1998, 2003). A simplified version of the activation mechanism that relied on recency and frequency of instances use was sufficient in capturing human choice behavior in several repeated binary-choice and probability-learning tasks (Lejarraga et al., 2010). According to this simplified mechanism, for each trial t, *activation* $A_{i,t}$ of instance i is:

$$\mathrm{A}_{i,t} = ln\left(\sum_{t_i \in \{1,\ldots,t-1\}} (t - t_i)^{-d}\right) + \sigma \times ln\left(\frac{1 - \gamma_{i,t}}{\gamma_{i,t}}\right) \tag{4}$$

Where d is a free decay parameter, and t_i is a previous trial when the instance i was created or its activation was reinforced due to an outcome observed in the task (the instance i is the one that has the observed outcome as the value in its outcome slot). The summation will include a number of terms that coincides with the number of times an outcome has been observed in previous trials and the corresponding instance i's activation that has been reinforced in memory (by encoding a timestamp of the trial t_i). Therefore, the activation of an instance corresponding to an observed outcome increases with the frequency of observation and with the recency of those observations. The decay parameter d affects the activation of an instance directly, as it captures the rate of forgetting. In ACT–R, the d parameter has a default value of 0.5 (Anderson & Lebiere, 1998, 2003). The higher the value of the d parameter, the faster the decay in memory, and the harder it is for the model to retrieve instances with outcomes that occurred many trials ago.

The $\gamma_{i,t}$ term is a random draw from a uniform distribution U (0, 1), and the $\sigma \times ln\left(\frac{1 - \gamma_{i,t}}{\gamma_{i,t}}\right)$ term represents Gaussian noise important for capturing the variability of human behavior. Again, σ is a free noise parameter that was defined above and that introduces noise in the retrieval of instances from memory.

Special treatment of the first trial. In the first trial, the model has no instances in its memory from which to calculate blended values. Therefore, the model makes a selection between instances that are pre-populated in its memory. Each of these instances corresponds to one of the decision options in a task with an outcome value pre-assigned to the instance. We used a value of +30 in the outcome slot of the two options' instances (Lejarraga et al., 2010). The +30 value is arbitrary, but importantly, it is greater than any possible outcomes in the task's problems and will trigger an initial exploration of the two options. In practice, the model makes a random selection for one of the two in the first trial. Because the +30 outcome in the two pre-populated instances is never actually observed in the task, they are never reinforced and their activation decays rapidly in memory. In the first trial, a decision is also based solely upon the blended values, and the inertia mechanism (Equation 1) is used from the second trial onwards.

The Stopping Point Rule: Accounting for Sample Size in the Sampling Paradigm

Although the IBL model above is the same across the two paradigms, the sampling paradigm involves a decision of when to stop sampling that is not necessary in the repeated-choice paradigm. To determine when the model stops sampling, a simple stopping rule is defined as a random draw from a distribution function.

The sampling paradigm's data from Hertwig et al. (2004) revealed that a negative-binomial distribution best fitted participants' sample-size: Negative-Binomial (2, 0.09) with $\chi^2(9) = 23.95$, $p < .01$. The distribution's plot is heavily right-tailed, which means most samples are small. The mean of the distribution is 21 samples of both options ($SD = 15$), and the median is 17. Similarly, the TPT's sampling paradigm's data gave the sample sizes for both options for 1,200 participants across the 60 problems in the estimation set (with 20 participants per problem; Erev, Ert, Roth, et al., 2010). A chi-square test revealed that a geometric distribution best fitted the participants' sample-size in TPT's estimation set: Geometric ($p = .082$) with $\chi^2(17) = 446.43$, $p < .001$. The distribution's plot is again heavily right-tailed in Erev, Ert, Roth, et al.'s (2010) data. The mean of the distribution is 11 samples of both options ($SD = 12$), and the median is 9 samples.

A random draw from these distribution functions only provides the number of samples, but the order in which the options are sampled in sampling and repeated-choice is determined by the same IBL model's mechanisms (Equations 1, 2, 3, and 4 above).

Quantitative Model Comparison: Sampling and Repeated-Choice

The comparison of models that make quantitative predictions regarding choices and cognitive processes in the different paradigms is expected to help form a comprehensive framework of DFE that is robust across tasks and paradigms. Using the same reasoning as task-specific models (Lejarraga et al., 2010), we attempt to demonstrate that our IBL model is not specific to one DFE paradigm.

The TPT (Erev, Ert, Roth, et al., 2010) was a commendable effort to advance current knowledge toward obtaining quantitative predictions to explain different observations in decisions from description and experience through model comparison. The goal of this model comparison effort, however, was to predict the results of specific experiments within each of three separate paradigms: decisions from description, sampling, and repeated-choice. Thus, models were submitted independently for each of the paradigms and were assessed by their ability to account for behavior in one of the two paradigms separately. Regardless of the behavioral similarities in the sampling and repeated-choice paradigms, very different models won the competition for each paradigms (Erev, Ert, Roth, et al., 2010). Furthermore, although the TPT followed a generalization procedure where models were calibrated in an estimation data set and then tested in a generalization (competition) data set, the TPT did not evaluate these models by their complexity. A model's ability to generalize and its complexity are two key properties of quantitative model comparisons and selection methods (Pitt et al., 2002). Finally, models in the TPT were evaluated for their accurate predictions in general choice behavior (i.e., R-rate) and not by process behavior (i.e., A-rate). Thus, none of the models submitted in the sampling paradigm were able to generate the sequence of samples that preceded choice. Similarly, none of the models submitted in the repeated-choice paradigm were evaluated by the accuracy of their prediction in sequential-choice behavior.

In the current research, we address all of the issues discussed above. We demonstrate that the same IBL model accounts for the choice and process behavior in both DFE paradigms. We compare the IBL model with the best models in both paradigms from the TPT (top five models in the sampling paradigm and top six models in the repeated-choice paradigm, a total of 11 models; Erev, Ert, Roth, et al., 2010). We account for model complexity by using the Akaike information criterion (AIC) that takes into account both a model's ability to predict human data and its complexity in terms of number of parameters contained (Pitt et al., 2002). We also expose all the competing models to multiple data sets for calibration and challenging generalizations, and we test the model's ability to predict both the general choice and sequential choice behavior at the same time.

Next, we summarize the top models in the sampling and the repeated-choice paradigms. We provide more detail for the winner models in the TPT, as they are the main comparison benchmarks to our IBL model.

Models in the Sampling Paradigm

1. The baseline model, called *primed sampler*, is a simplified version of the model by Erev, Glozman, and Hertwig (2008). The model assumes a simple rule: to take a random sample of k draws from each option (a free parameter with the best calibrated value in TPT's estimation set: $k = 5$) and to accumulate evidence (outcomes) based on its sampling of each option. The model selects the option with the highest accumulated evidence. In this model, all k samples drawn have equal weights in the accumulated value.

2. Another baseline model, called the *primed sampler with variability*, is a variation of the primed sampler model where the sample-size parameter k is varied between participants run on each problem. The sample size is uniformly drawn using one of the integers between 1 and k. Best calibrated value of the parameter was obtained with $k = 9$ in the TPT's estimation data set (Erev, Ert, Roth, et al., 2010). The model makes a choice between options exactly like the primed sampler model.

3. The runner-up model, called *sample by cumulative prospect theory and aspiration levels,* was submitted by Ahn and Picard (Erev, Ert, Roth, et al., 2010) and is based on prospect theory but is extended by reinforcement-learning algorithms. First, the prospect theory's weight and value functions are found based upon the current value of parameters. These weight and value functions are multiplied to derive an expected value for each option. These expected values are then used to compute a probability for each option using a Boltzmann distribution (Ahn, 2010). The model selects the option with the highest probability.

4. The second runner-up model, called *natural mean heuristic*, was submitted by Hau and Hertwig (Erev, Ert, Roth, et al., 2010) and is a variant of the original natural-mean heuristic model (Hertwig & Pleskac, 2008). This model selects the option with the larger average outcome during a sequential sampling process. The model takes between 1 and 20 samples from each option for each simulated participant. The probability of selecting the risky option is determined for each simulated participant based upon the number of samples (between 1 and 20) and a binomial distribution. Then, a weighted average of the probability of selecting the risky option over all 20 participants determines the final probability of selecting the risky option. To determine the weighted average, the model uses 20 weight parameters.

5. The winning model, called *Ensemble,* is discussed in Hau et al. (2008) and Erev, Ert, Roth, et al. (2010). This model assumes that each choice is made based on one of four equally likely rules: (1) a rule that is similar to the primed sampler model with variability, (2) another variant of the primed sampler model where the number of samples is drawn from a distribution of observed sample sizes in the TPT's estimation set, (3) a stochastic variant of the cumulative prospect theory, and (4) a stochastic version of the lexicographic priority heuristic adapted from the priority model proposed by Rieskamp (2008). The probability of selecting the risky option is determined by applying each of the four rules, and then the mean probability is derived and reported as the proportion of risky choices (Erev, Ert, Roth, et al., 2010). Although this model was the most successful in predicting behavior in the TPT's sampling paradigm, the accuracy of the Ensemble model came at a very high cost of its complexity (the model is made of four sub-models, one for each rule, and 40 different free parameters). As concluded by Erev, Ert, Roth, et al. (2010, p. 40): "the model is not easy to handle."

There are many other models in the sampling paradigm that did not participate or did not rank highly in the TPT. Different models in the sampling paradigm are explained and compared in Hau et al. (2008). For example, models based on heuristics include either some probability information (probability-based) or depend solely upon sampled outcomes (outcome-based). The popular outcome-based heuristic model include maximax and minimax (Luce &

Raiffa, 1957), which were originally proposed as models for decisions under ignorance; the natural-mean heuristic, which selects the option with the larger average outcomes during sampling (Hertwig & Pleskac, 2008); and the probability-based heuristic is the Lexicographic heuristic (Dawes, 1979), which selects the option with the highest and most frequently observed outcomes. Models of associative learning include the value-updating model (Hertwig, Barron, Weber, & Erev, 2006) and the fractional-adjustment model (March, 1996; Weber, Shafir, & Blais, 2004). In both models, learning involves changing the propensity to select an option based on the experienced outcomes (good experiences boost the propensity of selecting the associated choice, and bad experiences diminish it). In the value-updating model, participants update their estimates of the gamble's value after each new draw. Specifically, the model computes the weighted average of the previously estimated value and the value of the most recently experienced outcome for each option. The choice is made for the option with the higher of the two weighted-average values (Hertwig et al., 2006). Finally, a model based on cumulative prospect theory (called the two-stage model [TSM]) assumes that people first form subjective beliefs of the probability of outcomes, and then they enter these beliefs into cumulative prospect theory's weighting function, as suggested by Fox and Tversky (1998; Tversky & Fox, 1995). The TSM model first assesses the observed probability of non-zero outcomes in each option, calculates the expected value of each option, and chooses the option with the highest expected value (Hau et al., 2008).

An important conclusion from all the models described above is that none of the models in the sampling paradigm were designed to predict the sequence of sampling selections that humans make. All these models assumed some sampling procedure and then focus on predicting only the final consequential choice. The IBL model predicts not only the final consequential choice but also the sequence of sampling selections.

Models in the Repeated-Choice Paradigm

1. A baseline model, called *basic reinforcement learning* (Erev, Ert, Roth, et al., 2010), assumes a stochastic choice rule that resembles the Boltzmann distribution, where the temperature in the distribution is replaced by a weight. The weight value in trial $t + 1$ is a weighted average of the value and the obtained payoff in trial t (Erev & Barron, 2005; Erev, Bereby-Meyer, & Roth, 1999). The Boltzmann distribution provides a probability of selecting an option, and the model selects the option with the highest probability.

2. Another baseline model is the *normalized reinforcement learning* (NRL; Erev, Ert, Roth, et al., 2010). This model is an extension of the basic reinforcement learning model above with an additional assumption that the weight in the Boltzmann distribution is normalized by a payoff variability value. The payoff variability term is the weighted average of the difference between the obtained payoff at trial t and $t - 1$ (Erev & Barron, 2005; Erev et al., 1999). Again, the Boltzmann distribution provides a probability of selecting an option, and the model selects the option with the highest probability.

3. The best baseline model is the *explorative sampler (ES) with recency*. This model is a variation of the ES model (Erev, Ert, Roth, et al., 2010; Erev et al., 2008), which has been shown to outperform the reinforcement learning models and its variants (Erev, Ert, Roth, et al., 2010). The model can be summarized with the following assumptions (Erev et al., 2008):

(a) The agents are assumed to consider two cognitive strategies: exploration and exploitation. Exploration implies a random choice. The probability of exploration is 1 in the very first trial, and it reduces toward an asymptote with experience when information concerning the forgone payoffs is not available. The effect of experience on exploration depends on the expected length of the experiment (T). Exploration diminishes quickly when T is small and slowly when T is large.

(b) The experiences with each option include the set of observed outcomes yielded by this option in previous trials. When the payoffs are additionally limited to the obtained payoff, the subjective value of the very first outcome is recalled as an experience with all the options.

(c) Under exploitation, the agent draws (with replacement) a sample of m_t past experiences with each option. The first draw is the most recent experience with each option. All previous experiences are equally likely to be sampled in the remaining $m_t - 1$ draws. This assumption implies a hot stove effect: an increase in risk-aversion with experience.

(d) The value of m_t at trial t is assumed to be randomly selected from the set $\{1, 2, ., k\}$ where k is a free parameter.

(e) The recalled subjective values of the outcome x (from selecting option j) at trial t is assumed to be affected by two factors: regression to the mean of all the experiences with the relevant option (in the first $t - 1$ trials) and diminishing sensitivity. The estimated subjective value of each option is the mean of the subjective value of that option's sample in that trial. The model selects the option with the highest estimated subjective value.

4. The runner-up model, called *two-stage sampler*, is an improvement of the ES model (Ayal & Hochman, 2009). In addition to ES model's other assumptions, a two-stage sampler chooses an option based upon the difference between the average sampled value from history of observed outcomes and the low outcome. The average sampled value is computed just like in the ES model by repeated sampling for both options from history during exploitation. Like the ES model, the model selects the option with the highest average sampled value.

5. The second runner-up, the *NRL with inertia* (Erev, Ert, Roth, et al., 2010), is an extension of the NRL model. However, to determine the probability of each option, a stochastic choice rule is used. This rule resembles the Boltzmann distribution considering 68% contribution of the model's previous choice (i.e., inertia).

6. The winning model, *ACT–R with sequential dependencies and blending memory* (hereafter "ACT–R model"), was submitted by Stewart, West, and Lebiere (2009). This model is a variant of the IBL model presented in this article.

The ACT–R model assumes similarity-based inference, blending, and sequential dependencies. As an IBL model, experience in the ACT–R model is coded into an instance that includes the context, choice, and obtained outcome. The context in this model includes the two previous consecutive choices explicitly as part of the instance (Erev et al., 2008; Stewart et al., 2009). The retrieval of instances from memory considers all instances that are relevant for the current context and retrieves the one where the activation exceeds a threshold (captured by a free parameter, *retrieval threshold* [*RT*]). In addition to the *RT* parameter, this model includes the decay rate parameter and a noise parameter (d and σ in Equation 4). The choice is performed through the blending of instances in a way that the activation of each option is weighted by the probability of retrieving a corresponding instance from memory. The ACT–R model chooses an option with the highest blended value.

The ACT–R model and the IBL model reported here differ in several ways. These differences are already explained in Lejarraga et al. (2010) but are summarized here. The first difference is the context used in the ACT–R model and its absence in the IBL model. The IBL model only saves the previous choice and not the previous two choices. The second difference is in the constituting parts of the activation mechanism. In the IBL model, only the "base level" (frequency and recency) and noise are used (see Equation 4), whereas in the ACT–R model, a similarity mechanism was also used to match the current context with the context stored in memory. Also, the IBL model does not use the *RT* parameter (which limits the number of instances considered in the blending mechanism), whereas the ACT–R model uses the *RT* parameter as a free parameter. In the IBL model, all experiences have a probability (P_i) of being retrieved from memory and participate in blending, whereas only those instances with an activation greater than the *RT* are considered in blending for the ACT–R model. This distinction results in different types of instances that participate in the blending mechanism between the IBL and ACT–R models. Finally, the ACT–R model uses a different blending equation than the one used by our IBL model. The ACT–R model weighs the activation of each instance by the probability of retrieving an instance from memory (Lebiere, 1999), whereas the IBL model weighs the outcome of each instance by the same probability (Equation 2).

Again, there are other models in the repeated-choice paradigm that did not participate or did not rank at the top in the TPT. Reinforcement learning models are common in the repeated-choice literature. Many of these models assume that individuals are equipped with a set of strategies (Busemeyer & Myung, 1992; Erev & Barron, 2005). For example, Erev and Barron (2005) proposed a model to capture learning in DFE with partial-feedback and full-feedback that is based on reinforcement learning among cognitive strategies (RELACS). RELACS is a model that evolved from the proposed learning effects by Roth and Erev (1995; Erev & Roth, 1998). It assumes that the decision maker follows one of three cognitive strategies or decision rules in each decision encountered. The probability that any given strategy is used is determined by reinforcements derived from previous experiences using the strategy (Erev & Barron, 2005).

As summarized above, many models have attempted to capture the recent findings of DFE under each of the two experiential paradigms. These models were developed to account for particular characteristics in each, and they often carry different cognitive assumptions even within the same paradigm. To demonstrate the generality of the cognitive mechanisms proposed by IBLT, we compared the performance of the TPT's winners (the Ensemble model and the ACT–R model), the best baseline models, and the runner-up models in each of the two paradigms to the performance of one single IBL model in both paradigms.

SC Problems

Table 1 shows the SC problems that have been popularly used in DFE's repeated-choice (Barron & Erev, 2003) and sampling paradigms (Hertwig et al., 2004). Each problem has a High (H) and a Low (L) option based upon their expected values. For example, problem 1 shows that a "4" can occur with an 80% chance and a "0" with a 20% chance under the H option (with an expected value of 3.2); and for the L option, a "3" will occur with a 100% chance (with an expected value of 3). The last three columns in Table 1 present the proportion of H choices over all participants in the sampling (Pmax), repeated-choice (Pmax2), and description paradigms (Pmax) as reported in Hertwig et al. (2004) and in Barron and Erev (2003).

Hertwig et al. (2004) presented 100 participants with the SC problems of Table 1 in the sampling and description paradigms. Participants in the sampling paradigm (N = 50) saw two buttons on the computer screen and were told to sample freely before making a final choice. In each group, 25 participants were presented with three of the six problems, and the remaining 25 participants were presented with the other three problems. Barron and Erev (2003) used the same six problems in the repeated-choice paradigm. One hundred and forty-four students served as paid participants per problem. The experiment was run for four blocks

Table 1
The Proportion of Maximization (Pmax or Pmax2) in Humans Across Six Problems in Sampling, Repeated-Choice, and Description Paradigms

	Problem		Decisions from experience		Decisions from description
Number	High (H) option	Low (L) option	Sampling paradigm (from Hertwig et al., 2004) Pmax %[a]	Repeated paradigm (from Barron & Erev, 2003) Pmax2 %[b]	(from Hertwig et al., 2004) Pmax %[a]
1	4, 0.8; 0, 0.2	3, 1.0	88	62	36
2	4, 0.2; 0, 0.8	3, 0.25; 0, 0.75	44	48	64
3	−3, 1.0	−32, 0.1; 0, 0.9	28	41	64
4	−3, 1.0	−4, 0.8; 0, 0.2	56	61	28
5	32, 0.1; 0, 0.9	3, 1.0	20	24	48
6	32, 0.025; 0, 0.975	3, 0.25; 0, 0.75	12	30	64

[a] Derived based on data available from Hertwig et al. (2004). [b] Derived based on data provided by Ido Erev and reported in experiments in Barron and Erev (2003). The Pmax2 is the proportion of maximization in the second block of 100 trials (the problems were run for 4 blocks of 100 trials each, and Barron & Erev, 2003, reported proportion of maximization in the second block of 100 trials).

of 100 trials each, and groups of 24 participants were assigned to each of the six problems (1 through 6).

Calibration to the Proportion of Maximization (Pmax/Pmax2)

To evaluate the IBL model against the other models, the three models were calibrated to human data in the SC problems using the exact same procedures. The Ensemble model was calibrated to the sampling participants' data (Hertwig et al., 2004), the ACT–R model was calibrated to the repeated-choice participants' data (Barron & Erev, 2003), and the IBL model was calibrated to participant's data in both paradigms separately. Calibrating a model to participants' data means running the model in the same problems experienced by human participants to find the model's parameters values that minimize the squared distance between the model's Pmax (or Pmax2) and human's Pmax (or Pmax2). Appendix A explains the calculation of the mean squared deviation (MSD), correlation coefficient (*r*), and AIC, as well as the calibration procedure followed and the range in which different models' parameters were varied for calibration. Because we minimized the squared deviations, better AIC values are those that are more negative (see Appendix A for details).

Table 2 presents a summary of the calibrated parameters and the models' performance (in terms of AIC, MSD, and correlation coefficient *r*) in the two paradigms. The IBL model performs better than the Ensemble model (on Pmax) and the ACT–R model (on Pmax2). Figure 1A shows the Pmax for the final choice for human, IBL model, and Ensemble model in each of the six problems. Figure 1B shows the Pmax2 for the human, IBL model, and the ACT–R model across 100 trials in Block 2 (i.e., Trials 101–200). Both the IBL and ACT–R models explain the Pmax2 in human data very well; however, the Pmax2 predictions from the ACT–R model slightly overestimate the observed Pmax2 in human data.

The IBL model uses the same mechanisms to predict both the A-rate in the sampling and in the repeated-choice paradigm. For the sampling paradigm, we calculated the model's A-rate for the first 15 samples (the median number of samples in human data from Hertwig et al., 2004), and we also calculated the A-rate for the first 93 samples—which is the maximum number of samples in which there is least one participant in each of the IBL model and human data. For the repeated-choice paradigm, we calculated the A-rate for the first 200 trials (Blocks 1 and 2).

The results of the comparisons shown in Table 2 demonstrate that the IBL model's predictions worsen with a larger number of samples or trials used. Also, the model's A-rate predictions are slightly better in the repeated-choice than in the sampling paradigm. This result agrees with our expectations, because most participants rely on small samples and the number of participants decreases with greater number of samples in the sampling paradigm. Also, the predictions of the A-rate are better for the IBL model compared to the ACT–R model in the repeated-choice paradigm.

Figure 2A presents the A-rate for the IBL model and humans over the first 93 samples when averaged over all problems and participants in the sampling paradigm. The IBL model produces very good A-rate predictions for close to the median number of samples (= 15 samples). As the number of participants used to calculate the A-rate falls rapidly over increasing number of samples, the model's A-rate is more accurate for a smaller number of samples. Again, this is expected, given the variability that is introduced by a decreasing the number of participants in both the model and human data sets as the number of samples increases (the stopping point rule in the sampling paradigm).

Figure 2B presents the A-rate for the IBL model, ACT–R model, and human data over 199 trials in the repeated-choice paradigm. The IBL model predicts the A-rate in the human data very well; the ACT–R model overestimates the A-rate.

Overall, these results in the SC problems demonstrate that a single IBL model, and thus the same cognitive processes that it represents, can explain human behavior in two different DFE paradigms better than the TPT's winning models that were created for each paradigm separately. Furthermore, we demonstrate that the same model can capture the A-rate in both paradigms, where the only source of poor fit is the increased number of samples in the sampling paradigm, given that a variable and decreasing number of participants sample more.

However, these results are limited in several ways. First, the SC problems only provide a partial demonstration of what the IBL model can explain and predict. Given the limited number of problems and the similarities among them, there is a risk of over-fitting the human data. For example, the SC problems have no mixed outcomes (gains and losses together), and they have only low and high probability values and no medium probability values. Second, the results are only calibration to human data in these six problems. Because of this, it is possible that the IBL, the Ensemble, and the ACT–R models address the particular characteristics of the narrow set of problems but are unable to predict behavior in novel sets of problems. Third, although the models selected for comparison were the winners of the TPT in each of the two paradigms, it would be informative to compare the IBL model against other models to determine the characteristics that might be relevant to their performance in the DFE paradigms. These limitations are addressed in the next section.

The TPT

The same IBL model above was calibrated to the TPT's estimation set in the sampling and repeated-choice paradigms and then tested in the competition set. During both calibration and testing, the IBL model is compared to the winner, runner-up, and second runner-up models submitted in the TPT, as well as the baseline models (Erev, Ert, Roth, et al., 2010).

In recent research, the current IBL model (but without the inertia mechanism) was better than all other submitted models in the TPT's repeated-choice paradigm (Lejarraga et al., 2010). Here, there are three main extensions to that demonstration. First, we extend the demonstration across the DFE paradigms and compare the IBL model's performance to a larger number of models in each of the paradigms. Second, because the winning models and the IBL model differ in their number of parameters, we use more sophisticated model comparison techniques. In addition to the MSD and correlation coefficient *r*, we report the AIC measure from all models. Because we are interested in capturing the learning performance from different models, we report the MSD and *r* for a single trial in the sampling paradigm and across 100 trials in the repeated-choice paradigm. Our procedure is thus different from that used in Erev, Ert, Roth, et al. (2010) and Lejarraga et al.

Table 2

Summary of Calibration of IBL Model, Ensemble Model (Winner, Sampling Paradigm), and ACT–R Model (Winner, Repeated-Choice Paradigm) on the Proportion of Maximization (Pmax or Pmax2) in the SC Problems

AIC rank on Pmax	Model	Parameters	AIC (number of parameters, number of trials): Pmax	AIC: A-rate	MSD[a] (number of trials): Pmax	MSD: A-rate	Correlation (r)[a] (number of trials): Pmax	Correlation: A-rate
			Sampling paradigm					
1	IBL (calibrated upon the Pmax measure)	d = 0.29; σ = 0.27; *pInertia* = 0.22	**−2.6** (3, 1 final choice)	**−74.3** (3, first 15 samples)[b] **−339.3** (3, first 93 samples)[d]	**0.0002** (1 final choice)	**0.0047** (first 15 samples) **0.0243** (first 93 samples)	—[c]	**0.85** (first 15 samples) **0.46** (first 93 samples)
2	Ensemble (winner) (calibrated upon the Pmax measure)	*Alpha* = 1.63; *Beta* = 0.01; *Gamma* = 2.38; *Delta* = 0.49; *Lambda* = 0.16; *Mu* = 1.00; *Wmin* = 0.01; *Wp* = 0.91; *Sigma* = 1.00; *Throg* = 0.03; *Thrp* = 0.32; and 29 more parameters in Erev, Ert, Roth, et al. (2010)	**76.6** (40, 1)	cnd[e]	**0.0328** (1)	cnd[e]	—[c]	cnd[e]
			Repeated-choice paradigm					
1	IBL (calibrated upon the Pmax2 measure)	d = 0.86; σ = 0.58; *pInertia* = 0.48	**−627.4** (3, 100)	**−85.0** (first 15 trials) **−564.6** (first 93 samples) **−1,243.3** (first 200 trials)[f]	**0.0018** (100)	**0.0023** (first 15 trials) **0.0021** (first 93 trials) **0.0019** (first 200 trials)	**0.03** (100)	**0.90** (first 15 trials) **0.74** (first 93 trials) **0.68** (first 200 trials)
2	ACT–R (winner) (calibrated upon the Pmax2 measure)	d = 0.40; σ = 0.45; *RT* = 0.076	**−515.0** (3, 100)	**−43.5** (first 15 trials) **−280.8** (first 93 trials) **−602.7** (first 200 trials)	**0.0055** (100 trials)	**0.0368** (first 15 trials) **0.0458** (first 93 trials) **0.0477** (first 200 trials)	**0.11** (100 trials)	0.76 (first 15 trials) 0.56 (first 93 trials) 0.47 (first 200 trials)

Note. Values in bold indicate the AIC, MSD, and correlation r for the IBL and winner models. The Pmax and Pmax2 are evaluated over trials and are averaged over six problems and 25 participants per problem in the sampling paradigm, and over six problems and 24 participants per problem in the repeated-choice paradigm (there is a single trial of data in Pmax). IBL = the single computational model based on instance-based learning theory; ACT–R = Adaptive Control of Thought—Rational; SC = six choice; AIC = Akaike information criterion; MSD = mean squared deviation; *RT* = retrieval threshold.

[a] MSD and correlation coefficient (r) are calculated between the dependent measure (Pmax or Pmax2) from human data and that from model data for a single trial in the sampling paradigm (Pmax) and between trials 101 and 200 in the repeated-choice paradigm (Pmax2). [b] Fifteen is the median number of samples in human data. Thus, there are 75 human participants (median or half of 150) who sampled 15 or less times from both options. [c] The correlation coefficient (r) does not exist as there is only one trial. [d] The maximum number of samples in which there are at least one participant each in both the IBL model and human data in the sampling paradigm is 93. [e] Cannot be determined from the model as the model does not provide this information. [f] First two blocks in the repeated-choice paradigm (each block was 100 trials long and there were 4 blocks in total).

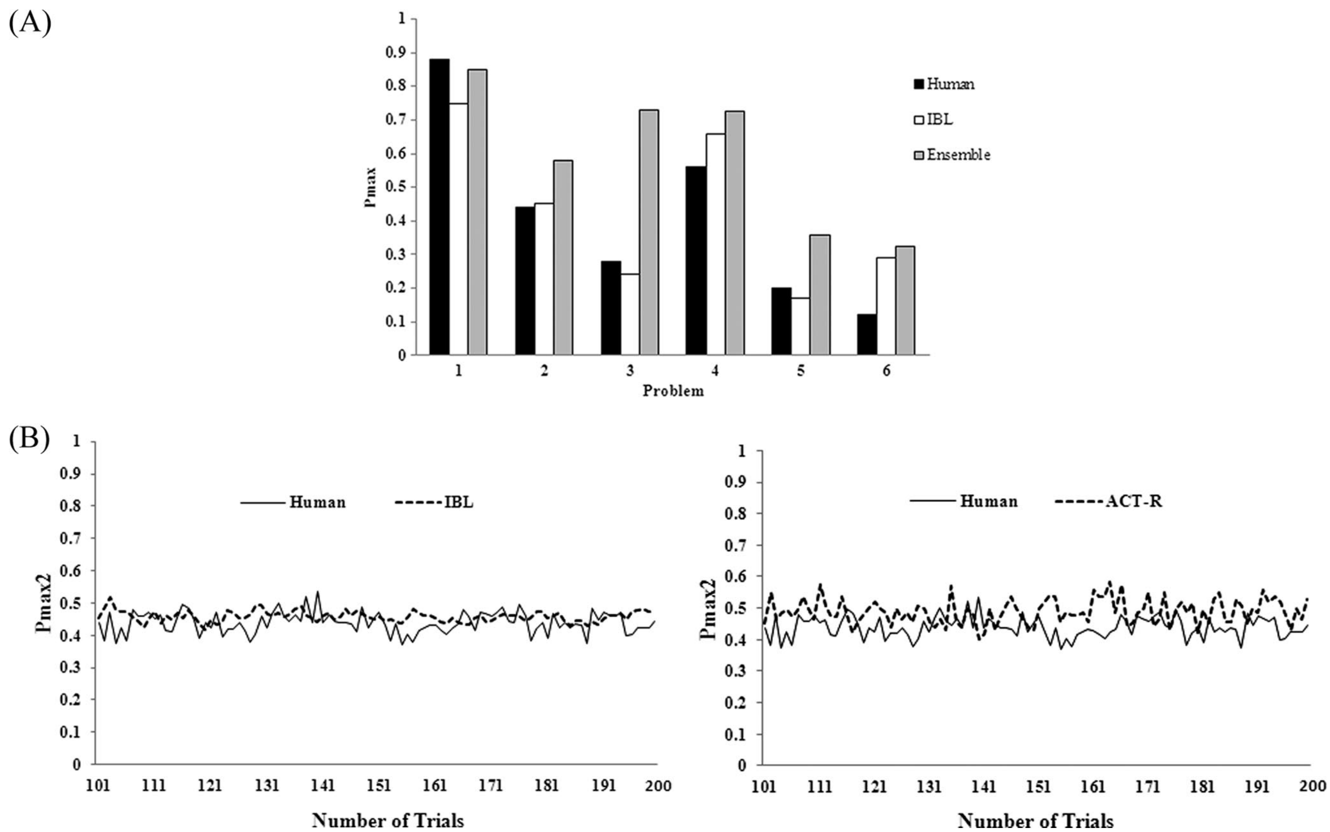

Figure 1. (A) Comparison of the proportion of maximization (Pmax) from the single computational model based on instance-based learning theory (IBL), Ensemble models, and human data in the six choice problems in the sampling paradigm. (B) The proportion of maximization in the second block (Pmax2) from the IBL model (left panel) and the Adaptive Control of Thought—Rational (ACT–R) model (right panel) compared to human data in the repeated-choice paradigm.

(2010), which report the MSD and *r* across the individual problems, not over trials. Third, in addition to demonstrating that the IBL model predicts the R-rate in both paradigms, we also present the predictions on the A-rate as an extension to the results reported in Erev, Ert, Roth, et al. (2010).

The TPT Data Sets

The TPT involved two data sets: an estimation set (60 problems) and a competition set (60 new problems derived using the same algorithm as the estimation set) in each of the DFE paradigms. The different models submitted to the TPT were calibrated to participants' choice in the estimation set. Later, these calibrated models were tested in the competition set with the parameters obtained in the estimation set, following the generalization method of model comparison (Busemeyer & Wang, 2000). Detailed information on methods and data collected in the TPT is explained in Erev, Ert, Roth, et al. (2010). In this article, we briefly summarize this information.

Different participants were tested in the sampling and repeated-choice paradigms. Each paradigm involved the same 60 problems in the estimation set and the same 60 problems in the competition set. All problems involved a choice between two unlabeled buttons, one associated with a safe option that offered a medium (M) outcome with certainty and the other associated with a risky option that offered a high (H) outcome with some probability (pH) and a low (L) outcome with the complementary probability (1-pH). Participants were not told which option was risky or safe and were simply asked to maximize their outcomes in each problem. The problem parameters M, H, pH, and L were generated randomly, and a selection algorithm ensured that the problems in each set were different in domain (positive, negative, and mixed outcomes) and probability values (high, medium, and low pH). The positive domain was such that each of the M, H, and L outcomes in a problem were positive numbers (>0). The mixed domain was such that one or two of the problem's outcomes were negative (<0). The negative domain was such that each of the problem's outcomes were negative numbers (<0). The low, medium, and high probability values corresponded to values of pH between .01 and .09, .1 and .9, and .91 and .99, respectively. The selection algorithm ensured that in both the estimation and competition sets there were 20 problems for each of the three domains and about 20 problems each for the three probability values. The resulting set of problems was thus large and representative of a broad diversity of problems.

In the repeated-choice paradigm's estimation and competition sets, 100 participants were randomly divided into five groups of 20

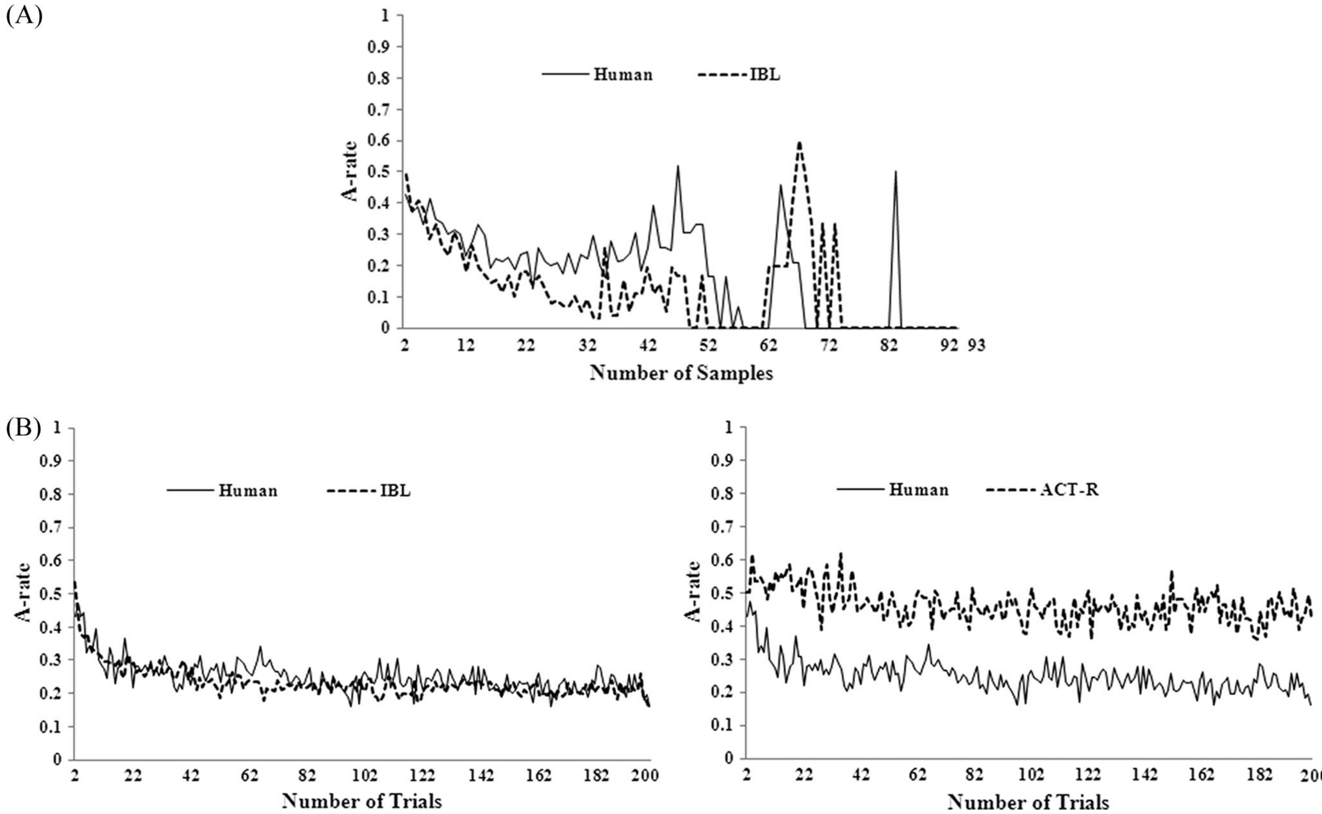

Figure 2. (A) The proportion of alternations (A-rate) for the single computational model based on instance-based learning theory (IBL) and humans in the sampling paradigm from the 2nd sample to the 93rd sample. (B) The A-rate for the IBL model (left panel) and the Adaptive Control of Thought—Rational (ACT–R) model (right panel) compared to human data between the 2nd and 200th trial in the repeated-choice paradigm.

participants each, and each group completed 12 out the 60 problems. The repeated-choice paradigm involved a block of 100 consequential trials per problem. For the sampling paradigm's estimation and competition sets, 40 participants were randomly assigned into two groups of 20 participants each, and each group completed 30 of the 60 problems.

Appendix B presents the experimental results in each paradigm and for each set. In this article, models submitted in the TPT are evaluated according to their ability to predict the proportion of risky choices (R-rate) in the competition set over 100 trials. The R-rate is averaged over 20 participants per problem and 60 problems in the repeated-choice paradigm, and it is averaged over 20 participants per problem and 60 problems for a final consequential choice in the sampling paradigm. In addition to the R-rate, we calculate the proportion of alternations (A-rate) in both paradigms over 100 trials (the A-rate in Trial 1 was assumed to be 0).

IBL Model's Performance in the Estimation Sets

We calibrated the IBL model in the estimation set separately for each paradigm, following the same procedures described in Appendix A.

Table 3 presents the summary of results from calibrating the IBL model in the estimation set, and the results from the best and baseline models (Erev, Ert, Roth, et al., 2010). In the sampling paradigm, the IBL model performed better than the Ensemble (winner), runner-up, and second runner-up models on the AIC measure. Interestingly, the IBL model performed better than the Ensemble model, even without accounting for the model complexity. The IBL model's AIC was slightly higher than the corresponding values in the baseline primed sampler and the primed sampler with variability models. This observation is due to the higher number of parameters in the IBL model, as its MSD value is clearly lower than both the baseline models. An advantage of the IBL model over all the other models is that it provides reasonably good predictions for the A-rate, whereas none of the competing models can produce those predictions in the sampling paradigm. The IBL model's predictions for the A-rate are derived using the first 86 samples, which is the maximum number of samples where there is at least one participant sampling in both the IBL model and human data.

In the repeated-choice paradigm, the IBL model performed better than the ACT–R model (winner) and the NRL with inertia (second runner-up) for both the R-rate and A-rate; however, the AIC value of IBL is slightly higher than the remaining models.

IBL Model's Generalization to the Competition Sets

The IBL model was evaluated against human data in the competition set in each of the two paradigms, using the parameters

Table 3
Summary of Calibration of the IBL Model and Other Models in the TPT's Estimation Set

AIC rank on R-rate	Model	Parameters	AIC (number of parameters, number of trials) R-rate	AIC (number of parameters, number of trials) A-rate	MSD (number of trials) R-rate	MSD (number of trials) A-rate	Correlation (r) (number of trials) R-rate	Correlation (r) (number of trials) A-rate
			Sampling paradigm					
3	IBL (calibrated upon the R-rate measure)	$d = 8.79$; $\sigma = 1.17$; $pInertia = 0.63$	**−2.8** (3, 1)	**−32.1** (3, first 9 samples)[a] **−227.4** (3, first 86 samples)[c]	**0.0002** (1)	**0.0145** (first 9 samples) **0.0663** (first 86 samples)	—[b]	**0.96** (first 9 samples) **−0.05** (first 86 samples)
6	Ensemble (winner)	40 parameters (refer to Erev, Ert, Roth, et al., 2010)	**72.3** (40, 1)	cnd[d]	**0.0005** (1)	cnd[d]	—[b]	cnd[d]
4	Sample by CPT and aspiration levels (runner-up)	$x = 2.07$; $y = 1.31$; $z = 0.71$; $v = 7.53$; $r = 12.64$; $m = 0.02$	2.7 (6, 1)	cnd[d]	0.0001 (1)	cnd[d]	—[b]	cnd[d]
5	Natural mean heuristic (second runner-up)	A vector of 20 weights that determined the probability of different samples (refer to Erev, Ert, Roth, et al., 2010)	33.5 (20, 1)	cnd[d]	0.0015 (1)	cnd[d]	—[b]	cnd[d]
2	Primed sampler (interesting baseline)	$k = 5$	−4.0 (1, 1)	cnd[d]	0.0024 (1)	cnd[d]	—[b]	cnd[d]
1	Primed sampler with variability (best baseline)	$k = 9$	−4.1 (1, 1)	cnd[d]	0.0022 (1)	cnd[d]	—[b]	cnd[d]
			Repeated-choice paradigm					
5	IBL (calibrated upon the R-rate measure)	$d = 5.27$; $\sigma = 1.46$; $pInertia = 0.09$	**−705.6** (3, 100)	**−321.0** (3, 100)	**0.0008** (100)	**0.0380** (100)	**0.92** (100)	**0.86** (100)
6	ACT–R (winner)	$d = 5.00$; $\sigma = 0.35$; $RT = -1.60$	**−617.9** (3, 100)	**−280.1** (3, 100)	**0.0020** (100)	**0.0572** (100)	**0.92** (100)	**0.69** (100)
3	Two-stage sampler (runner-up)	$w = 1.00$; $d = 0.55$; $b = 0.10$; $f = 0.3$; $\kappa = 6.00$; $\epsilon = 0.11$; $r = .01$	−760.5 (7, 100)	−366.1 (7, 100)	0.0004 (100)	0.0224 (100)	0.94 (100)	0.88 (100)
7	NRL with inertia (second runner-up)	$w = 0.14$; $\lambda = 1.05$; $q = 0.32$; $i = 0.50$	−309.9 (4, 100)	−309.7 (4, 100)	0.0416 (100)	0.0417 (100)	0.93 (100)	0.76 (100)
1	Basic reinforcement learning (interesting baseline − 1)	$\omega = 0.15$; $\lambda = 1.00$	−772.0 (2, 100)	−358.3 (2, 100)	0.0004 (100)	0.0267 (100)	0.95 (100)	0.84 (100)
4	NRL (interesting baseline − 2)	$w = 0.15$; $\lambda = 1.10$	−744.7 (2, 100)	−298.8 (2, 100)	0.0006 (100)	0.0484 (100)	0.83 (100)	0.72 (100)
2	Explorative sampler with recency (best baseline)	$\beta = 0.10$; $w = 0.30$; $\epsilon = 0.12$; $k = 8$	−763.0 (4, 100)	−353.9 (4, 100)	0.0004 (100)	0.0268 (100)	0.95 (100)	0.90 (100)

Note. Values in bold indicate the AIC, MSD, and correlation r for the IBL and winner models. IBL = the single computational model based on instance-based learning theory; TPT = Technion Prediction Tournament; AIC = Akaike information criterion; MSD = mean squared deviation; CPT = cumulative prospect theory; ACT–R = Adaptive Control of Thought—Rational; RT = retrieval threshold; NRL = normalized reinforcement learning.
[a] Nine is the median number of samples in human data. [b] The correlation coefficient (r) does not exist as there is a single trial. [c] The maximum number of samples in which there are at least one participant each in both the IBL model and human data is 86. [d] Cannot be determined from the model as the model does not provide this information.

obtained in the corresponding paradigms' estimation set. Table 4 presents the summary of the generalization results and the AIC values for the best and baseline models reported in Erev, Ert, Roth, et al. (2010).

In the sampling paradigm, the IBL model performed better than all other competing models, but it also performed slightly below the two baseline models on AIC. The baseline models have an advantage over the IBL model in the number of parameters, but the IBL model has an advantage over all the existent models in the sampling paradigm as it is able to generate sampling sequences and predict an A-rate with good accuracy.

In the repeated-choice paradigm, the IBL model performs as well as the ACT–R model on the R-rate, better than all the other competing models, and slightly below the NRL baseline model on AIC and MSD.

Figure 3A shows the A-rate from the IBL model and human data over the first 98 samples (from 2nd sample to 99th sample) in the sampling paradigm. The 99th sample is the maximum number of samples for which there is at least a single participant sampling in both model and human data. The IBL model shows more accurate predictions of the A-rate for about the first 9 samples (median). With larger number of samples, the number of participants in both data sets drops, and the predictions deviate from the human data. Figure 3B show the A-rate predictions from the IBL and the ACT–R model compared to the A-rate in human data in the competition set. Both models seem to overestimate the A-rate in human data. However, the A-rate predictions from IBL model seem to converge over trials toward the human A-rate curve, whereas the predictions from the ACT–R model seem to diverge away over trials.

IBL Model's Generalization Across the Sampling and Repeated-Choice Paradigms

For each of the estimation and competition sets, the IBL model calibrated to the human data was evaluated against human data in each paradigm, using the parameters obtained during calibration. The only difference between the models used in the sampling and repeated-choice paradigms was the stopping rule: When we generalized the repeated-choice parameters to the sampling data set, we used the stopping rule of the sampling paradigm; and when we generalized the sampling parameters to the repeated-choice data set, we removed the stopping rule, as the number of trials in this case is fixed to 100.

The results, summarized in Table 5, demonstrate that the sole difference between the two is the stopping rule. The same model with the same parameters calibrated to one of the paradigms is able to generalize very accurately to the data in the other paradigm.

A Challenging Generalization: From the TPT's Estimation Set to SC Problems

This section extends the generalization procedure and evaluates the IBL model against other TPT models in SC problems using the parameters derived from calibration in the TPT's estimation set. Because the TPT problems and the SC problems are not identical in their structure and they belong to different experiments with different populations, we expect this to be a more challenging generalization for all models involved. We argue that the TPT's generalization condition had the characteristics of traditional cross-validation, rather than generalization (see also Gonzalez, Dutt, & Lejarraga, 2011, for a similar argument). The problems in the competition set were similar to the problems in the estimation set because they had the same structure, were created with the same algorithm, and involved a similar participant population: There was no difference between the R-rate in the estimation set (49%) and the R-rate in the competition set (44%) (U = 1,573.5; $Z = -1.195$, *ns*) for the sampling paradigm. Similarly, there was no difference between the R-rate in the estimation set (40%) and the R-rate in the competition set (38%) (U = 1,666.5; $Z = -0.701$, *ns*) for the repeated-choice paradigm.

To run all the models created for the TPT in the SC problems, we made a simple assumption: When a SC problem has both H and L probabilistic options (Problems 2 and 6; see Table 1), we treat the L option as a medium option by replacing the two outcomes in the L option with their expected value. This assumption is the simplest change one needs to make to run all TPT models without changing their structure or working.

Table 6 presents the generalization results of the IBL model and other TPT models, when these models are generalized in the SC problems using calibrated parameters derived in the TPT's estimation set. In the sampling paradigm, the IBL model's AIC value for Pmax was better than all other models' except for the best baseline model (primed sampler with variability), where the IBL model's AIC was slightly higher. In the repeated-choice paradigm, the IBL model had the best AIC for Pmax2 and A-rate compared to other models.

Figure 4A shows the IBL and Ensemble models' Pmax predictions compared to the human data's Pmax in the SC problems. Figure 4B shows the IBL and ACT–R models' Pmax2 predictions compared to the Pmax2 in human data. Although both models provide good predictions, the IBL model's predictions are better than the ACT–R model's.

Figure 5A shows the A-rate predictions from the IBL model compared to the human data. The model's A-rate predictions are reasonably good for the first 15 samples (median). Figure 5B shows the A-rates from IBL (left panel) and ACT–R (right panel) models compared to the human data. Both models produce reasonable but slightly different generalizations. The AIC for the A-rate puts the IBL model above the winner and all other models in the repeated-choice paradigm.

Reconciling the Sampling and Repeated-Choice Paradigms

Human risk, maximization behavior, and sequential selection of options in both DFE paradigms can be explained by the same learning processes and mechanisms in IBLT. These involve the storage and retrieval of situation-decision-outcome instances governed by inertia, blending, frequency, recency, and noisy memory retrieval processes. The only difference between the two paradigms is the addition of a stopping rule in the model to account for each individual's sampling size in the sampling paradigm.

The IBL model involved three parameters: *pInertia*, decay (d), and noise (σ). We observe similarity of the d and σ parameter fit values across paradigms and within each data set. In the SC problems, the calibration of the parameters resulted in $d = 0.29$, and $\sigma = 0.27$ for the sampling paradigm, and $d = 0.86$, and $\sigma =$

Table 4

Summary of Generalization of the IBL Model and Other Models in the TPT's Competition Set (the Parameters of IBL Model Were Obtained in the Estimation Set in Table 3 by Calibrating the Model Using the R-Rate Measure in the Two Paradigms)

AIC rank on R-rate	Model	AIC (number of parameters, number of trials) R-rate	AIC A-rate	MSD (number of trials) R-rate	MSD A-rate	Correlation (r) (number of trials) R-rate	Correlation A-rate
			Sampling paradigm				
2	IBL	**−3.4** (3, 1)	**−41.3** (3, first 9 samples)[a] **−382.1** (3, first 99 samples)	**0.0001** (1)	**0.0051** (first 9 samples) **0.0198** (first 99 samples)	—[b]	**0.98** (first 9 samples) **0.32** (first 99 samples)
5	Ensemble winner)	**73.2** (40, 1)	cnd[c]	**0.0011** (1)	cnd[c]	—[b]	cnd[c]
3	Sample by CPT and aspiration levels runner-up)	−1.4 (6, 1)	cnd[c]	0.0000 (1)	cnd[c]	—[b]	cnd[c]
4	Natural mean heuristic second runner-up)	31.6 (20, 1)	cnd[c]	0.0015	cnd[c]	—[b]	cnd[c]
1	Primed sampler interesting baseline)	−6.0 (1, 1)	cnd[c]	0.0004 (1)	cnd[c]	—[b]	cnd[c]
1	Primed sampler with variability best baseline)	−6.0 (1, 1)	cnd[c]	0.0003 (1)	cnd[c]	—[b]	cnd[c]
			Repeated-choice paradigm				
2	IBL	**−750.7** (3, 100)	**−313.8** (3, 100)	**0.0005** (100)	**0.0408** (100)	**0.96** (100)	**0.86** (100)
2	ACT−R winner)	**−750.7** (3, 100)	**−290.5** (3, 100)	**0.0005** (100)	**0.0515** (100)	**0.96** (100)	**0.75** (100)
4	Two-stage sampler runner-up)	−668.0 (7, 100)	−370.5 (7, 100)	0.0011 (100)	0.0214 (100)	0.93 (100)	0.88 (100)
6	NRL with inertia second runner-up)	−353.2 (4, 100)	−317.9 (4, 100)	0.0270 (100)	0.0384 (100)	0.95 (100)	0.78 (100)
5	Basic reinforcement learning interesting baseline − 1)	−588.3 (2, 100)	−253.3 (2, 100)	0.0027 (100)	0.0763 (100)	0.97 (100)	0.56 (100)
1	NRL interesting baseline − 2)	−792.6 (2, 100)	−304.2 (2, 100)	0.0004 (100)	0.0459 (100)	0.97 (100)	0.74 (100)
3	Explorative sampler with recency best baseline)	−688.1 (4, 100)	−356.1 (4, 100)	0.0009 (100)	0.0262 (100)	0.89 (100)	0.85 (100)

Note. Values in bold indicate the AIC, MSD, and correlation r for the IBL and winner models. IBL = the single computational model based on instance-based learning theory; TPT = Technion Prediction Tournament; AIC = Akaike information criterion; MSD = mean squared deviation; CPT = cumulative prospect theory; ACT–R = Adaptive Control of Thought—Rational; NRL = normalized reinforcement learning.

[a] Nine is the median number of samples in human data. [b] The correlation coefficient (r) does not exist as there is a single trial. [c] Cannot be determined from the model.

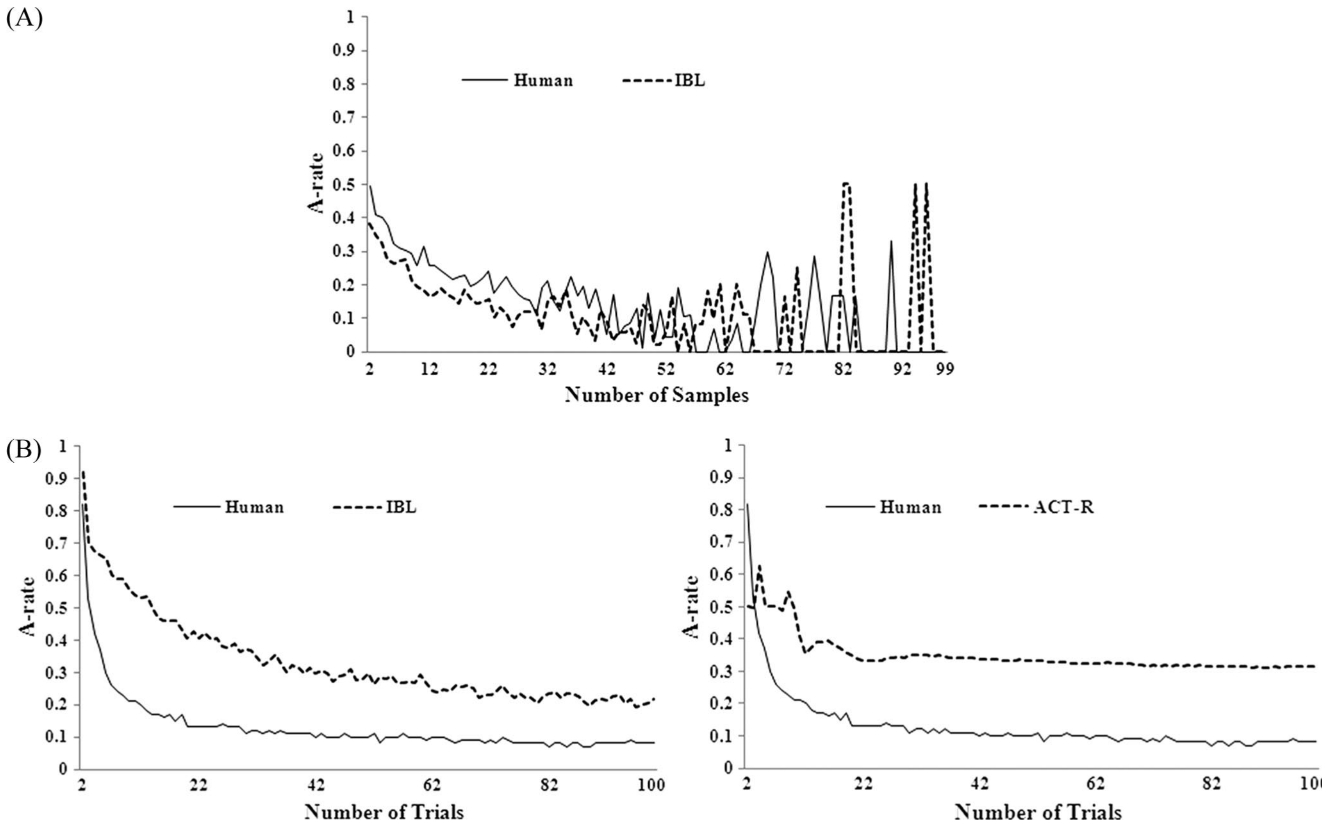

Figure 3. (A) The proportion of alternations (A-rate) for the single computational model based on instance-based learning theory (IBL) and human data in the sampling paradigm from the 2nd sample to the 99th sample (the 99th sample is the maximum number of samples for which there is one participant each in the model and human data, respectively) of the competition data set. (B) The A-rate for the IBL and the Adaptive Control of Thought—Rational (ACT–R) model and human data between the 2nd and 100th trial in the repeated-choice paradigm of the competition data set.

0.58 for the repeated-choice paradigm. The d and σ values found are within the range of default or common values in the ACT–R architecture (Anderson & Lebiere, 1998; Wong et al., 2010). The parameters are also similar for the sampling and repeated-choice paradigms within the TPT data sets. In the TPT's estimation set, the calibration resulted in $d = 8.79$, and $\sigma = 1.17$ for the sampling paradigm, and $d = 5.27$, and $\sigma = 1.46$ for the repeated-choice paradigm. This supports the similarity of the cognitive processes involved in two DFE paradigms.

In contrast to the similar values of those parameters *within* each data set, there is a marked difference *between* the SC problems and TPT data sets. The d and σ values in the TPT estimation set are much higher than those in the SC problems. This observation suggests that recency and noise of memory retrieval depends on the characteristics and diversity of the problems being represented. The problems in the SC set are less diverse than the problems in the TPT sets. As discussed before, the TPT problems include a larger range of probability values, domains, and outcomes. In addition, problems in the TPT involve outcomes expressed with decimals, whereas the SC problems only involve integer numbers. Less recency and noise are needed to fit the human behavior in the SC problems, as they are less diverse than the problems in the TPT. The low values of the d parameter in the SC problems indicate slower decay of memory and easier recall of distant outcomes encoded in memory instances. Outcomes in the SC problems are easier to remember and recall than in the TPT problems. As the diversity of problems increases in the TPT, higher values of d and σ are necessary. Faster decay of memory and more noisy retrieval processes requires greater reliance on recent experiences of outcomes encoded in instances.

Most importantly, this observation applies equally to both paradigms. Thus, the leading explanation of the description–experience gap in the repeated-choice paradigm of "reliance on recent experiences" is common in both paradigms, and it depends on the characteristics and diversity of the problems confronted. Rare outcomes would be harder to recall as the diversity of the problems confronted increases. Consequently, the contribution of a rare outcome's activation to the blended value will be small, and thus the model will behave as if these rare outcomes have less impact than they deserve according to their objective probability.

Also, the alternation effects over trials were very similar in both paradigms. This can be observed in Figures 2, 3, and 5 and most clearly in the generalization demonstration of Table 5. In both paradigms, the A-rate decreases over an increased number of samples or trials. Furthermore, the same IBL model calibrated in one paradigm with the same parameters predicts the A-rate of the

Table 5
Summary of Generalization of the IBL Model Calibrated in the Repeated-Choice/Sampling Paradigm to the Sampling/Repeated-Choice Paradigm, in Each of the Estimation and Competition Data Sets

Model	Parameters	AIC (number of parameters, number of trials)		MSD (number of trials)		Correlation (r) (number of trials)	
		R-rate	A-rate	R-rate	A-rate	R-rate	A-rate
			Estimation set				
Repeated-choice to sampling	$d = 5.27$; $\sigma = 1.46$; *pInertia* = 0.09	−1.51 (3,1)	−18.1 (3, first 9 samples) −208.1 (3, first 86 samples)	0.00054 (1)	0.0687 (first 9 samples) 0.0828 (first 86 samples)	—[a]	0.95 (first 9 samples) 0.05 (first 86 samples)
Sampling to repeated-choice	$d = 8.79$; $\sigma = 1.17$; *pInertia* = 0.63	−794.1 (3, 100)	−551.8 (3, 100)	0.00033 (100)	0.0037 (100)	0.93 (100)	0.90 (100)
			Competition set				
Repeated-choice to sampling	$d = 5.27$; $\sigma = 1.46$; *pInertia* = 0.09	−3.21 (3,1)	−15.7 (3, first 9 samples) −177.2 (3, first 82 samples)	0.0001 (1)	0.0896 (first 9 samples) 0.1071 (first 86 samples)	—[a]	0.98 (first 9 samples) 0.47 (first 82 samples)
Sampling to repeated-choice	$d = 8.79$; $\sigma = 1.17$; *pInertia* = 0.63	−535.6 (3, 100)	−577.9 (3, 100)	0.0016 (100)	0.0029 (100)	0.96 (100)	0.90 (100)

Note. The same parameters obtained from the corresponding paradigm were used in the evaluation of the model in the opposite paradigm. IBL = the single computational model based on instance-based learning theory; AIC = Akaike information criterion; MSD = mean squared deviation.
[a] The correlation coefficient (r) does not exist as there is a single trial.

other paradigm. The only difference is the added stopping rule needed in the sampling paradigm. Thus, results suggest that in both paradigms, humans move gradually from the exploration of options to their exploitation using the same cognitive mechanisms for the sequential selection of alternatives. However, the *pInertia* parameter cannot explain this similar transition. In the SC problems, the calibration of the parameters resulted in *pInertia* = 0.22 for the sampling paradigm and *pInertia* = 0.48 for the repeated-choice paradigm, whereas the TPT's estimation set resulted in *pInertia* = 0.63 in the sampling paradigm and *pInertia* = 0.09 in the repeated-choice paradigm.

There are several possible explanations for these results. First, because the parameters of the IBL model were calibrated to the R-rate and not the A-rate, different values of *pInertia* may represent the model's tradeoffs between handling the decreasing A-rate and maintaining the similar R-rate over time. Future work should investigate the tradeoffs existent between the R- and A-rates in both paradigms. Second, because we used a static inertia rule that does not change across trials and participants, the different values of *pInertia* may just represent the best compromise across trials to fit human behavior, but it cannot explain the transition from exploration to exploitation. More recent accounts of the exploration–exploitation tradeoffs in the repeated-choice paradigm suggest that inertia is a function of surprising outcomes (Nevo & Erev, 2011). The probability of inertia is not static but rather decreases when the recent outcomes are surprising (Erev, Ert, & Roth, 2010; Nevo & Erev, 2011). Future work should investigate inertia as a function of surprise across DFE paradigms.

Main Lessons From Model Comparison

Finding a single model that can explain and generalize in diverse conditions and paradigms of DFE is an important challenge (Pitt & Myung, 2002; Pitt et al., 2002). As we have observed in this research, there are several good models that are able to predict human choice behavior quite accurately within each paradigm. In fact, many of the models within each paradigm are slightly different quantifications of similar psychological assumptions: learning from a finite number of experiences, giving more weight to more recent experiences, and accounting for the frequency of experiences. However, the small advantage the IBL model has over other models within each paradigm (according to the AIC R-rate/Pmax measures) should not obscure the more important benefits. First, the IBL model demonstrated that behavior in the sampling and repeated-choice paradigms is equivalent at the general (R-rate) level and at the sequential process (A-rate) level. It remains a challenge for other models to predict these similarities across the two paradigms with their own same mechanisms. Second, we now provide an important guideline for modelers who would attempt such challenge: The only difference between the two paradigms is the stopping rule. Third, the IBL model is the only currently existent model that can predict the sequence of sampling selections that humans make. Other models in the sampling paradigm assume a sampling procedure (often a random rule) and focus on predicting only the final consequential choice. Fourth, the IBL model makes consistently better predictions than the winning models of the TPT competition.

Table 6
Summary of Generalization of the IBL Model and Other Models in the Six Problems (the Parameters of All Models Were Obtained in the TPT's Estimation Set in Table 3)

		AIC (number of parameters, number of trials)		MSD (number of trials)		Correlation (r) (number of trials)	
AIC rank on Pmax	Model	Pmax	A-rate	Pmax	A-rate	Pmax	A-rate
		Sampling paradigm					
2	IBL	**−2.6** (3,1)	**−201.1** (3,first 62 samples)[a] **−64.2** (3, first 15 samples)[c]	**0.0002** (1)	**0.0354** (first 62 samples) **0.0103** (first 15 samples)	—[b]	**0.35** (first 62 samples) **0.78** (first 15 samples)
5	Ensemble (winner)	**69.3** (40,1)	cnd[d]	**0.0000** (1)	cnd[d]	—[b]	cnd[d]
3	Sample by CPT and aspiration levels (runner-up)	8.0 (6,1)	cnd[d]	0.0181 (1)	cnd[d]	—[b]	cnd[d]
4	Natural mean heuristic (second runner-up)	36.0 (20,1)	cnd[d]	0.0180 (1)	cnd[d]	—[b]	cnd[d]
2	Primed sampler (interesting baseline)	−2.6 (1,1)	cnd[d]	0.0092 (1)	cnd[d]	—[b]	cnd[d]
1	Primed sampler with variability (best baseline)	−3.4 (1,1)	cnd[d]	0.0044 (1)	cnd[d]	—[b]	cnd[d]

		AIC (number of parameters, number of trials)		MSD (number of trials)		Correlation (r) (number of trials)	
AIC rank on Pmax2	Model	Pmax2	A-rate	Pmax2	A-rate	Pmax2	A-rate
		Repeated-choice paradigm					
1	IBL	**−632.6** (3,100)	**−975.8** (3,100)	**0.0017** (100)	**0.0074** (100)	**−0.07** (100)	**0.73** (100)
2	ACT–R (winner)	**−592.1** (3,100)	**−840.4** (3,100)	**0.0025** (100)	**0.0282** (100)	**−0.07** (100)	**0.57** (100)
4	Two-stage sampler (runner-up)	−570.3 (7,100)	−921.0 (7,100)	0.0029 (100)	0.0093 (100)	0.05 (100)	0.69 (100)
7	NRL with inertia (second runner-up)	−223.8 (4,100)	−756.7 (4,100)	0.0985 (100)	0.0219 (100)	−0.02 (100)	0.71 (100)
6	Basic reinforcement learning (interesting baseline − 1)	−491.9 (2,100)	−852.0 (2,100)	0.0070 (100)	0.0138 (100)	−0.08 (100)	0.71 (100)
5	NRL (interesting baseline − 2)	−532.0 (2,100)	−734.9 (2,100)	0.0047 (100)	0.0249 (100)	−0.11 (100)	0.67 (100)
3	Explorative sampler with recency (best baseline)	−582.0 (4,100)	−871.5 (4,100)	0.0027 (100)	0.0124 (100)	0.05 (100)	0.63 (100)

Note. Values in bold indicate the AIC, MSD, and correlation r for the IBL and winner models. IBL = the single computational model based on instance-based learning theory; TPT = Technion Prediction Tournament; AIC = Akaike information criterion; MSD = mean squared deviation; CPT = cumulative prospect theory; ACT–R = Adaptive Control of Thought—Rational; NRL = normalized reinforcement learning.
[a] The maximum number of samples in which there are at least one participant each in both the IBL model and human data is 62. [b] The correlation coefficient (r) does not exist as there is a single trial. [c] Fifteen is the median number of samples in human data. [d] Cannot be determined from the model.

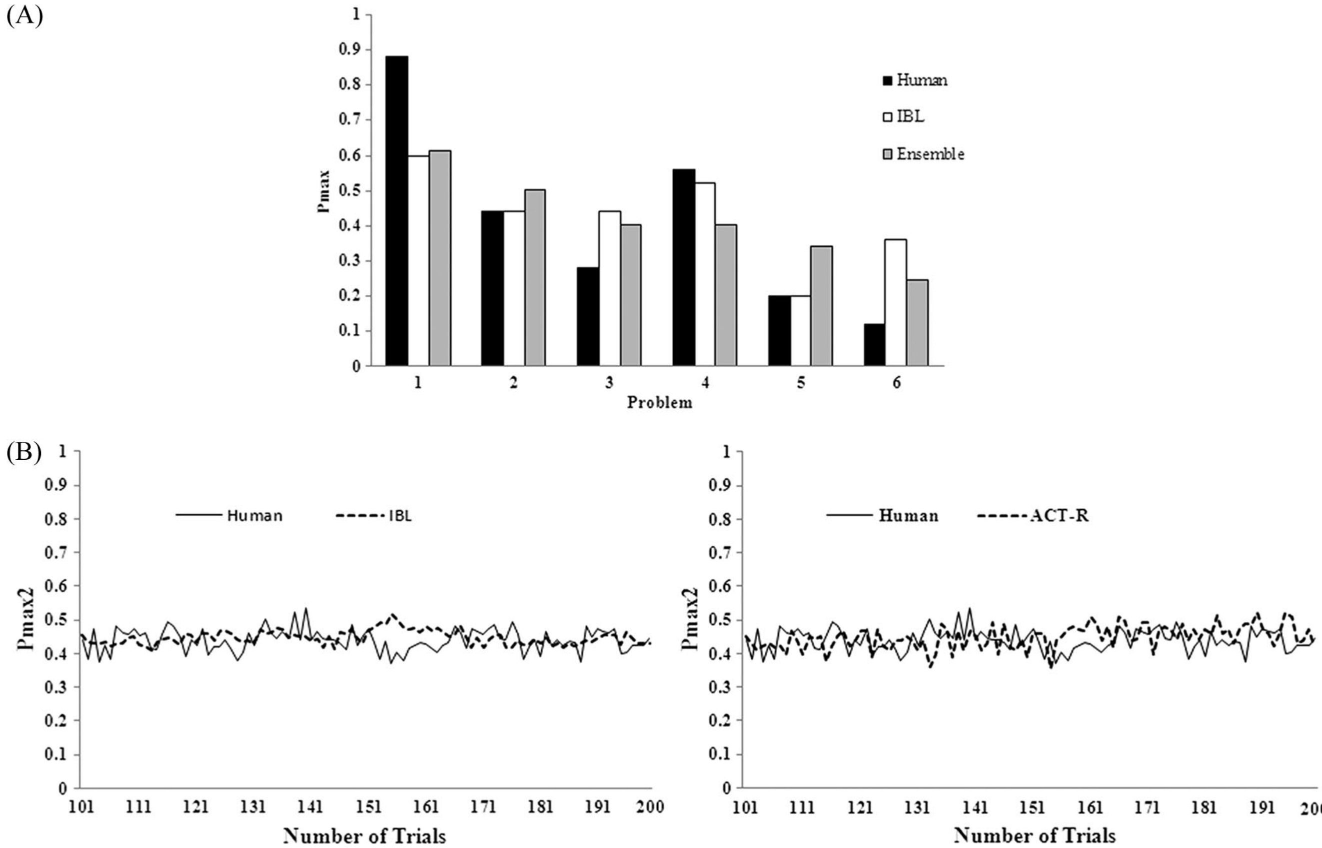

Figure 4. (A) The proportion of maximization (Pmax) predictions from the single computational model based on instance-based learning theory (IBL) and Ensemble models and that observed in human data in the generalization to the six choice problems in the sampling paradigm. (B) The proportion of maximization in the second block (Pmax2) predictions from the IBL model, Adaptive Control of Thought—Rational (ACT–R) model, and that observed in human data in the repeated-choice paradigm.

An important lesson from model comparison is the tradeoffs between the accuracy of models' explanations and the complexity needed to make those predictions (Pitt et al., 2002). Figure 6 demonstrates the tradeoffs in the quality of predictions of different models in their fitting and generalization in different data sets, and the complexity present in these models. To visualize these tradeoffs, we normalized values of the two parts of the AIC measure computed for the R-rate/Pmax/Pmax2. The y-axis plots the normalized value of the AIC's $t * \ln\frac{SSE}{t}$ term, which refers to the accuracy in capturing human data. Furthermore, the x-axis plots the normalized value of AIC's $2 * k$ term, which refers to a model's complexity in terms of the number of parameters (k is the number of parameters in a model). The optimal point in all these graphs is the (0, 0): a model with no parameters that fits and predicts human data perfectly. Of course such a model does not exist: All models are wrong, and they are only approximations of representations of human behavior. However, the distance from the origin represents how close a model is to the theoretical best model. The IBL model is not the absolute winner across all calibrations and generalizations when one accounts for both dimensions of different models. However, being the only model that is common across the sampling and repeated-choice paradigms, the IBL model ranks highly across all the calibration and generalization processes. In fact, it ranks as the best model in the challenging exercise of generalizing from the TPT's estimation set to the SC problems for both paradigms.

There are many limitations to the model comparison process used here. A main disadvantage comparing models that differ considerably in their assumptions and underlying mathematical processes is that qualitative comparisons are very difficult to make. We have been able to describe the differences between the winning ACT–R model in the TPT and the IBL model qualitatively (see above and Lejarraga et al., 2010) because both are IBL models and both rely on similar ACT–R mechanisms (Anderson & Lebiere, 1998). We have discussed the differences in the instance representation, the use of inertia, and the *RT* parameter. However, this comparison would be practically impossible to make between the Ensemble and IBL models. The mechanisms, the processes, and the assumptions are not comparable between the two. Thus, relying on the quantitative predictions that the models can make on the same tasks has many advantages, but limits qualitative understanding of the commonalities shared by different models.

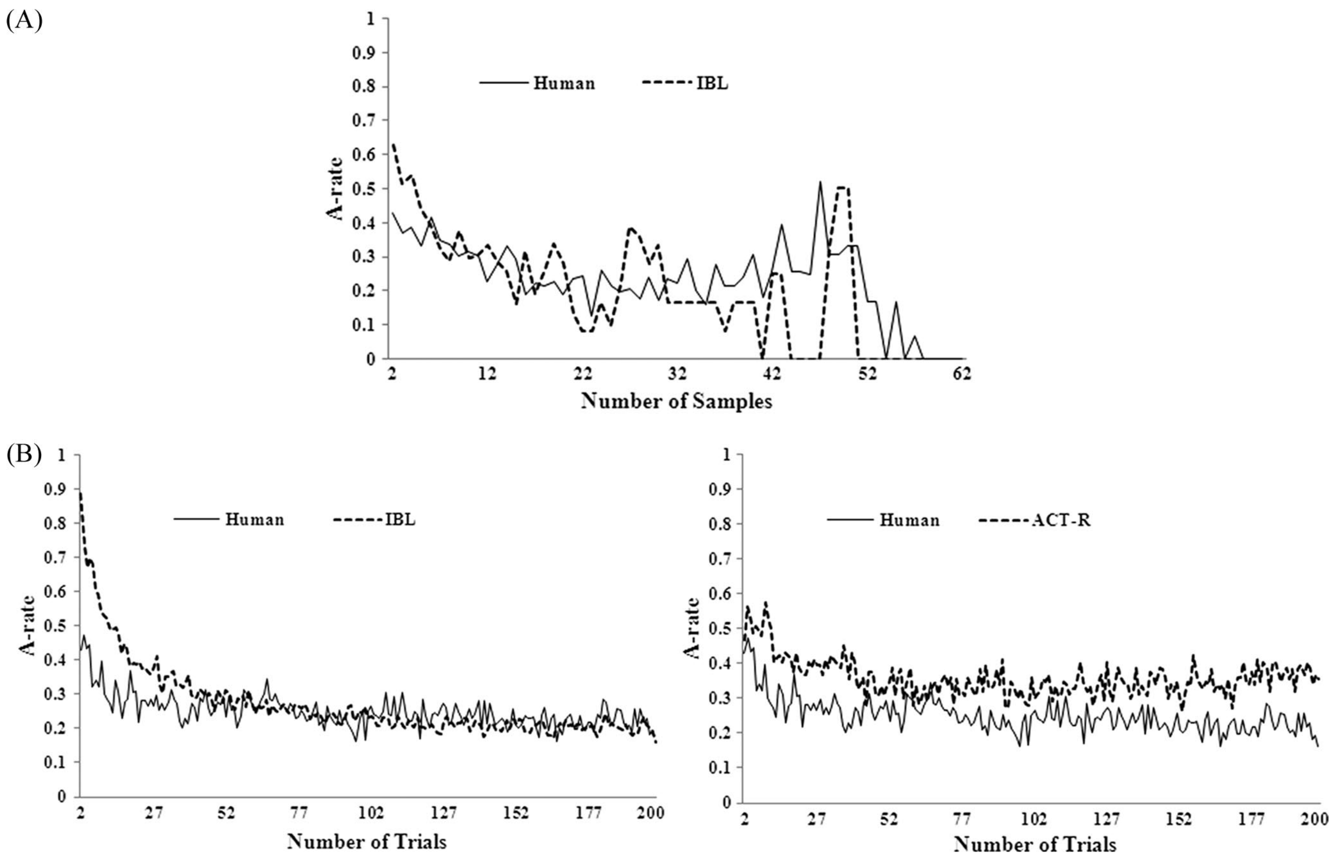

Figure 5. (A) The proportion of alternations (A-rate) for the single computational model based on instance-based learning theory (IBL) and humans in the sampling paradigm from the 2nd sample to the 62nd sample. (B) The A-rate for the IBL model, Adaptive Control of Thought—Rational (ACT–R) model, and human data between the 2nd and 200th trial in the repeated-choice paradigm.

Boundaries and Extensions of the IBL Model to Dynamic Decision-Making Tasks

Despite how advantageous the simple IBL model is in choice tasks, the current model is not expected to generalize to more complex situations. In fact, the IBL model shown here is only one instantiation of IBLT that is robust across many DFE choice tasks. As the complexity of the tasks increases, so will the need for additional mechanisms proposed by the original activation equation (Anderson & Lebiere, 1998), the constraints and mechanisms described in IBLT (Gonzalez et al., 2003), and perhaps new unknown mechanisms.

Complexity in dynamic tasks is defined not only by an increased number of options and more attributes in an instance but also, and most importantly, by the complexity created by the interactions of these elements over time, called *dynamic complexity* (Sterman, 2000). Researchers have found that decision makers remain suboptimal even in the simplest dynamic system after repeated practice, unlimited time, and performance incentives (Diehl & Sterman, 1995; Paich & Sterman, 1993; Sterman, 1989a, 1989b). A common cause is the multiple feedback processes, time delays, and nonlinearities involved in these systems and the “inability” to deal with such complexity (Cronin & Gonzalez, 2007; Cronin, Gonzalez, & Sterman, 2009). However, we need to understand and explain the underlying cognitive mechanisms leading to the learning difficulties in dynamic tasks (Gonzalez et al., 2003). Next, we discuss some ways in which the IBL model might be extended from the least dynamic tasks presented in this article to the most dynamic tasks in Edward’s (1962) taxonomy.

The simple IBL model for binary-choice tasks is easy to expand to tasks with more than two options and to tasks involving more than one

---

Figure 6 (opposite). The normalized value of Akaike information criterion’s (AIC’s) $t * \ln\frac{SSE}{t}$ term versus the normalized value of AIC’s $2 * k$ term for different models in the different calibration and generalization tests. The point (0, 0) is theoretically the best point: a model with no parameters that fits the human data perfectly. The AIC is calculated upon the R-rate/Pmax/Pmax2. SC = six choice; IBL = the single computational model based on instance-based learning theory; TPT = Technion Prediction Tournament.

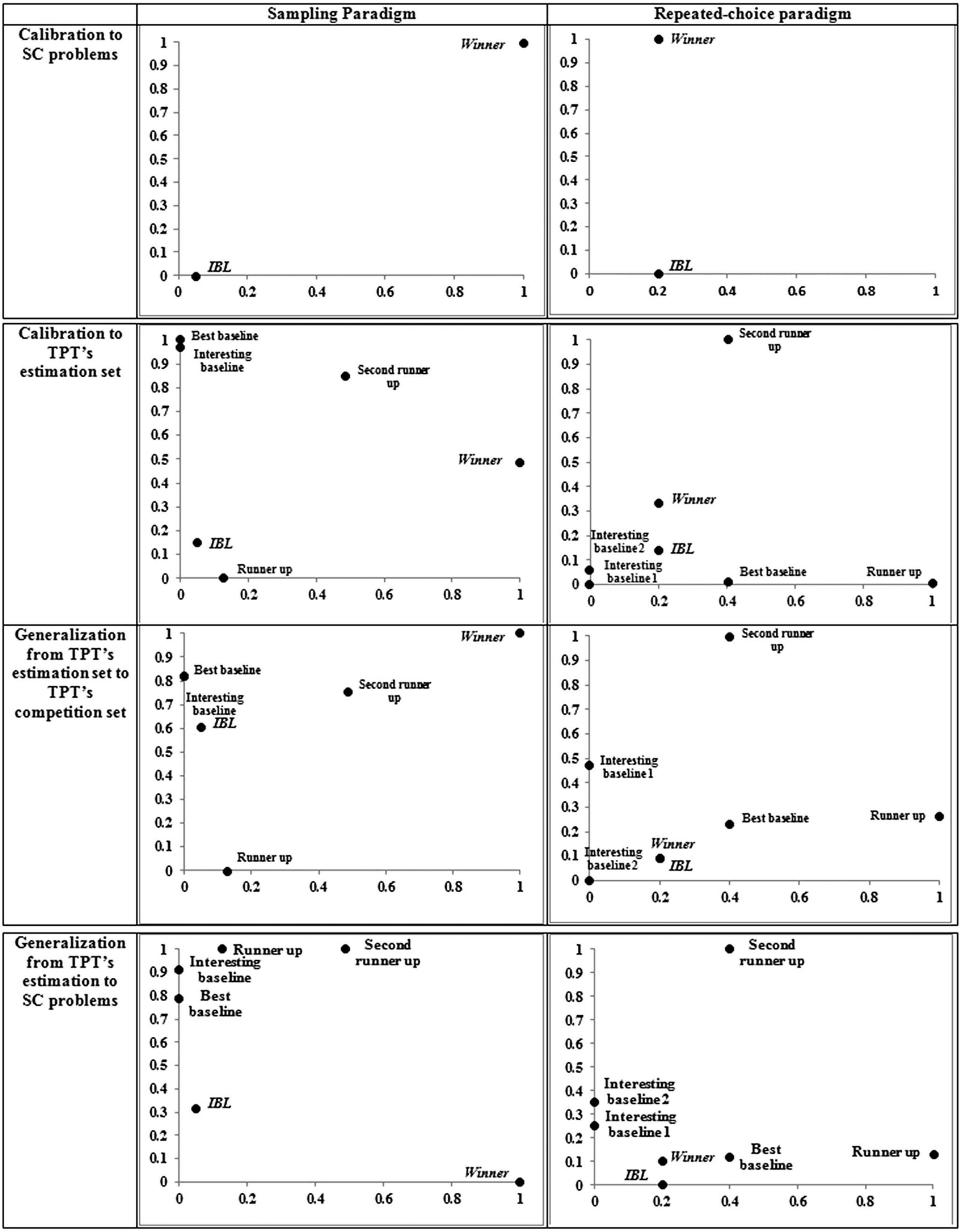

Figure 6 (opposite).

player. In fact, an example of this extended version of the IBL model recently received second place in the Market Entry Prediction Competition (Gonzalez, Dutt, & Lejarraga, 2011) organized by Erev, Ert, and Roth. Given that the choice rule entails selecting the option with the highest blended value and the model expands to account for any number of blended values (Equation 2), it is able to explain choice among more than two options. We have also developed an IBL model for the Iowa Gambling Task, which involves four options (Dutt & Gonzalez, 2011), and more generally are extending IBLT to account for behavior in social interactions (Gonzalez & Martin, 2011).

A more challenging extension of the IBL model will involve more dynamic choice tasks. For example, a complex binary-choice task where the probability of an outcome in trial t depends upon the nature of observed outcomes in trials $t - 1$ and $t - 2$, similar to the Markov process described by Biele et al. (2009). In the simple binary-choice task, the attributes that define the options are simple and irrelevant: two blank buttons on the screen, and both appear the same. As the attributes that define the options become more relevant and conceptually and contextually rich, other mechanisms such as spreading activation (Anderson & Lebiere, 1998; Gonzalez et al., 2003) will help determine the priorities and weights of those attributes in instances. This characteristic offers interesting opportunities to investigate the synergies between decision field theory (Busemeyer & Townsend, 1993) and IBLT. The decision field theory would help explain how the attributes of an instance are built over time. IBLT (and, in general, the ACT–R theory) assumes a prior definition of an instance structure.

Other mechanisms such as similarity through partial matching (Anderson & Lebiere, 1998; Gonzalez et al., 2003) will be needed because situation attributes can change over time. In this case, a similarity function will need to be defined to capture the relationship between the attributes in memory and those currently in the dynamic task. As explained in IBLT (Gonzalez et al., 2003), a defined threshold is needed to determine when an instance is "similar enough" to a situation. This characteristic will directly influence the retrieval and reuse of instances in memory.

Finally, there are currently many questions left to be answered regarding the processing of feedback, particularly when it is delayed. Thus, we expect that new extensions of the current IBL model will be needed for these situations. In IBLT, feedback is generally associated with the outcome of an executed decision. This is not problematic when each decision leads to one immediate outcome, but it is problematic when multiple decisions lead to one delayed outcome. In the model developed by Gonzalez et al. (2003) for a dynamic task, it was assumed that the feedback would apply equally to all the decisions made between the last time feedback was received and the current time. Although that appeared to be a reasonable assumption, it is not clear whether this assumption applies more generally to other tasks. Behavioral studies are needed to understand what factors mediate the effects of aggregated and delayed feedback, and how this feedback is processed in different decision conditions and decision interdependencies (Gonzalez, 2005).

Summary and Implications for the Psychology of Decision Making

Results from our quantitative model comparisons suggest that to account for behavior in DFE more generally, models can be more generic and still provide robust explanations of human behavior. Human behavior in both paradigms can be explained by the learning processes proposed by IBLT. The sampling and repeated-choice processes, the sequential choice behavior, and the overall risk-taking behavior can all be explained by the storage and retrieval of instances governed by the inertia, blending, frequency, and recency mechanisms from IBLT. The only significant difference in behavior between the two paradigms is the stopping point, which is needed in the sampling and not in the repeated-choice tasks.

In addition to being an encompassing model for both DFE paradigms, the IBL model addresses many limitations of current models that have been designed independently for the two paradigms. The IBL model is able to account for both the general maximization behavior and the alternation behavior *at the same time*. Influential models of repeated-choice have found a very weak relationship between the proportion of maximization and sequential dependencies (Erev & Barron, 2005), and they have captured the overall human choices without a good understanding of sequential dependencies (Erev & Barron, 2005; Rapoport, Erev, Abraham, & Olson, 1997). The IBL model generates accurate predictions of the alternation rates in both the sampling and repeated-choice paradigms while also producing accurate predictions of the maximization rates. Thus, because the IBL model is able to produce a sequence of choices (or samples), the blended values of the previous t choices or samples would predict the $t + 1$ choice or sample. The model's ability to predict the A-rate in both data sets led to the conclusion that a gradual transition from exploration of options to their exploitation is very similar in both paradigms. The learning curves of A-rate show a decrease that is similar in both paradigms, and the predictions of the IBL model are very accurate in both paradigms when we account for the number of samples (the stopping point). Thus, we would expect that if people in the repeated-choice paradigm were asked to stop making decisions whenever they felt comfortable making their last choice like in the sampling paradigm, both paradigms would become equivalent. Similarly, we would expect that if people in the sampling paradigm were forced to sample a fixed number of times like in the repeated-choice paradigm, both paradigms would become equivalent.

Although simple, the two paradigms represent two major forms in which we make DFE in the real-world. Both paradigms are important and relevant to understanding natural learning processes and decision making in situations where the options, outcomes, and probabilities are unknown. Decisions can often be made after sampling options that have no real consequences, but there are also decisions that need to be made with no opportunity to sample. The fact that both types of DFE can be explained by the same cognitive mechanisms and that the only difference between the two paradigms is the stopping rule has significant implications for training and preparing people for costly or consequential decisions. For example, our results suggest that the use of simulators that allow people to sample and learn the options without real consequences may be just as effective as making consequential decisions. Furthermore, in some circumstances, sampling may promote the acquisition, durability, and transferability of decision-making skills before confronting situations with real consequences.

References

Ahn, H. (2010). *Modeling and analysis of affective influences on human experience, prediction, decision making, and behavior* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge. Retrieved from http://affect.media.mit.edu/publications.php

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

Anderson, J. R., & Lebiere, C. (2003). The Newell test for a theory of mind. *Behavioral and Brain Sciences, 26,* 587–601. doi:10.1017/S0140525X0300013X

Ayal, S., & Guy, H. (2009). Ignorance or integration: The cognitive process underlying choice behavior. *Journal of Behavioral Decision Making, 22,* 455–474. doi:10.1002/bdm.642

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making, 16,* 215–233. doi:10.1002/bdm.443

Biele, G., Erev, I., & Ert, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *Journal of Mathematical Psychology, 53,* 155–167. doi:10.1016/j.jmp.2008.05.006

Busemeyer, J. R., & Diederich, A. (2009). *Cognitive modeling.* New York, NY: Sage.

Busemeyer, J. R., & Myung, I. J. (1992). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology: General, 121,* 177–194. doi:10.1037/0096-3445.121.2.177

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432–459. doi:10.1037/0033-295X.100.3.432

Busemeyer, J. R., & Wang, Y. M. (2000). Model comparison and model selections based on generalization criterion methodology. *Journal of Mathematical Psychology, 44,* 171–189. doi:10.1006/jmps.1999.1282

Camilleri, A. R., & Newell, B. R. (2011). When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychonomic Bulletin & Review, 18,* 377–384. doi:10.3758/s13423-010-0040-2

Cronin, M., & Gonzalez, C. (2007). Understanding the building blocks of system dynamics. *System Dynamics Review, 23,* 1–17. doi:10.1002/sdr.356

Cronin, M., Gonzalez, C., & Sterman, J. D. (2009). Why don't well-educated adults understand accumulation? A challenge to researchers, educators and citizens. *Organizational Behavior and Human Decision Processes, 108,* 116–130. doi:10.1016/j.obhdp.2008.03.003

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34,* 571–582. doi:10.1037/0003-066X.34.7.571

Denrell, J., & March, J. G. (2001). Adaption as information restriction: The hot stove effect. *Organization Science, 12,* 523–538. doi:10.1287/orsc.12.5.523.10092

Diehl, E., & Sterman, J. D. (1995). Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes, 62,* 198–215. doi:10.1006/obhd.1995.1043

Dutt, V., & Gonzalez, C. (2011). *Making instance-based learning theory usable and understandable: The Instance-Based Learning Tool.* Manuscript submitted for publication.

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—ICCM 2009* (pp. 77–82). Manchester, England: University of Manchester.

Dutt, V., Yamaguchi, M., Gonzalez, C., & Proctor, R. W. (2010). *Instance-based learning models of SRC and Simon effects.* Manuscript submitted for publication.

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,* 912–931. doi:10.1037/0033-295X.112.4.912

Erev, I., Bereby-Meyer, Y., & Roth, A. E. (1999). The effect of adding a constant to all payoffs: Experimental investigation, and implications for reinforcement learning models. *Journal of Economic Behavior & Organizations, 39,* 111–128. doi:10.1016/S0167-2681(99)00028-1

Erev, I., Ert, E., & Roth, A. E. (2010). A choice prediction competition for market entry games: An introduction. *Games, 1,* 117–136. doi:10.3390/g1020117

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., . . . Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making, 23,* 15–47. doi:10.1002/bdm.683

Erev, I., Glozman, I., & Hertwig, R. (2008). What impacts the impact of rare events. *Journal of Risk and Uncertainty, 36,* 153–177. doi:10.1007/s11166-008-9035-z

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review, 88,* 848–881. doi:10.2307/117009

Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review, 107,* 659–676. doi:10.1037/0033-295X.107.4.659

Fox, C. R., & Tversky, A. (1998). A belief-based account of decision under uncertainty. *Management Science, 44,* 879–895. doi:10.1287/mnsc.44.7.879

Gonzalez, C. (2005). Decision support for real-time dynamic decision making tasks. *Organizational Behavior and Human Decision Processes, 96,* 142–154. doi:10.1016/j.obhdp.2004.11.002

Gonzalez, C., Best, B. J., Healy, A. F., Kole, J. A., & Bourne, L. E., Jr. (2011). A cognitive modeling account of simultaneous learning and fatigue effects. *Journal of Cognitive Systems Research, 12,* 19–32. doi:10.1016/j.cogsys.2010.06.004

Gonzalez, C., & Dutt, V. (2010). Instance-based learning models of training. In *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 2319–2312). San Francisco, CA: Human Factors and Ergonomics Society.

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,* 136–162. doi:10.3390/g2010136

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, NY: Palgrave Macmillan.

Gonzalez, C., Lerch, F. J., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science, 27,* 591–635. doi:10.1207/s15516709cog2704_2

Gonzalez, C., & Martin, J. M. (2011). Scaling up instance-based learning theory to account for social interaction. *Negotiation and Conflict Management Research, 4,* 110–128. doi:10.1111/j.1750-4716.2011.00075.x

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description–experience gap in risky choice: The role of sample size and experienced probabilities. *Journal of Behavioral Decision Making, 21,* 493–518. doi:10.1002/bdm.598

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,* 534–539. doi:10.1111/j.0956-7976.2004.00715.x

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2006). Decisions from experience: Sampling and updating of information. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 72–91). New York, NY: Cambridge University Press.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences, 13,* 517–523. doi:10.1016/j.tics.2009.09.004

Hertwig, R., & Pleskac, T. J. (2008). The game of life: How small samples render choice simpler. In N. Chater & M. Oaksford (Eds.), *The probabilistic mind: Prospects for rational models of cognition* (pp. 209–236). Oxford, England: Oxford University Press.

Hills, T. T., & Hertwig, R. (2010). Information search in decisions from experience: Do our patterns of sampling foreshadow our decisions? *Psychological Science, 21,* 1787–1792. doi:10.1177/0956797610387443

Holland, J. H. (1975). *Adaptation in natural and artificial systems.* Ann Arbor, MI: University of Michigan Press.

Jakobsen, T. (2010). Genetic algorithms. Retrieved from http://subsimple.com/genealgo.asp

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–291. doi:10.2307/1914185

Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlations. *Psychological Review, 107,* 397–402. doi:10.1037/0033-295X.107.2.397

Lebiere, C. (1999). *Blending: An ACT–R mechanism for aggregate retrievals.* Paper presented at the 6th Annual ACT–R Workshop at George Mason University, Fairfax County, VA.

Lebiere, C., Gonzalez, C., & Martin, M. (2007). Instance-based decision making model of repeated binary choice. In *Proceedings of the 8th International Conference on Cognitive Modeling* (pp. 67–72). Ann Arbor, MI: Psychology Press.

Lejarraga, T., Dutt, V., & Gonzalez, C. (2010). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making.* Advance online publication. doi:10.1002/bdm.722

Luce, R. D., & Raiffa, H. (1957). *Games and decisions.* New York, NY: Wiley.

March, J. G. (1996). Learning to be risk averse. *Psychological Review, 103,* 309–319. doi:10.1037/0033-295X.103.2.309

Martin, M. K., Gonzalez, C., & Lebiere, C. (2004). Learning to make decisions in dynamic environments: ACT–R plays the beer game. In *Proceedings of the Sixth International Conference on Cognitive Modeling* (pp. 178–183). Pittsburgh, PA: Carnegie Mellon University/University of Pittsburgh.

Nevo, I., & Erev, I. (2011). *On surprise, change, and the effect of recent outcomes.* Manuscript submitted for publication.

Paich, M., & Sterman, J. D. (1993). Boom, bust, and failures to learn in experimental markets. *Management Science, 39,* 1439–1458. doi:10.1287/mnsc.39.12.1439

Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences, 6,* 421–425. doi:10.1016/S1364-6613(02)01964-2

Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review, 109,* 472–491. doi:10.1037/0033-295X.109.3.472

Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making, 23,* 1–14. doi:10.1002/bdm.681

Rapoport, A. (1975). Research paradigms for studying dynamic decision behavior. In D. Wendt & C. Vlek (Eds.), *Utility, probability, and human decision making* (Vol. 11, pp. 349–375). Dordrecht, the Netherlands: Reidel.

Rapoport, A., Erev, I., Abraham, E. V., & Olson, D. E. (1997). Randomization and adaptive learning in a simplified poker game. *Organizational Behavior and Human Decision Processes, 69,* 31–49. doi:10.1006/obhd.1996.2670

Rieskamp, J. (2008). The probabilistic nature of preferential choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1446–1465. doi:10.1037/a0013646

Roth, A. E., & Erev, I. (1995). Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior, 8,* 164–212. doi:10.1016/S0899-8256(05)80020-X

Sterman, J. D. (1989a). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes, 43,* 301–335. doi:10.1016/0749-5978(89)90041-1

Sterman, J. D. (1989b). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science, 35,* 321–339. doi:10.1287/mnsc.35.3.321

Sterman, J. D. (2000). Learning in and about complex systems. *Reflections: The SoL Journal, 1,* 24–51. doi:10.1162/152417300570050

Stewart, T. C., West, R., & Lebiere, C. (2009). Applying cognitive architectures to decision making: How cognitive theory and the equivalence measure triumphed in the Technion Prediction Tournament. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the Thirty-First Annual Conference of the Cognitive Science Society* (pp. 561–566). Amsterdam, Netherlands: Cognitive Science Society.

Tversky, A., & Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological Review, 102,* 269–283. doi:10.1037/0033-295X.102.2.269

Weber, E. U., Shafir, S., & Blais, A. R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation. *Psychological Review, 111,* 430–445. doi:10.1037/0033-295X.111.2.430

Wong, T. J., Cokely, E. T., & Schooler, L. J. (2010). An online database of ACT–R parameters: Toward a transparent community-based approach to model development. In D. D. Salvucci & G. Gunzelmann (Eds.), *Proceedings of the 10th International Conference on Cognitive Modeling* (pp. 282–286). Philadelphia, PA: Drexel University.

(*Appendices follow*)

Appendix A

Estimating the Parameters for the Single Computational Model Based on Instance-Based Learning Theory (IBL), the Ensemble Model, and the Adaptive Control of Thought—Rational (ACT–R) Model

Definition of Mean Squared Deviation (MSD)

$$\text{MSD} = \frac{\sum_{i=1}^{t}(x_{model,i} - x_{human,i})^2}{n}$$

Where $x_{model,i}$ and $x_{human,i}$ refer to the dependent measure (e.g., Pmax, Pmax2) in the model and human data for trial i. The dependent measure has been averaged over all participants and problems for each trial. The t is the total number of trials in each of the model and human data sets. The smaller the MSD value, the better model's prediction about human data.

Definition of Correlation Coefficient (*r*)

$$r = \frac{t * \sum_{i=1}^{t} x_{model,i} * x_{human,i} - \sum_{i=1}^{t} x_{model,i} * \sum_{i=1}^{t} x_{human,i}}{\sqrt{t * \sum_{i=1}^{t} x_{model,i}^{2} - \left(\sum_{i=1}^{t} x_{model,i}\right)^2} * \sqrt{t * \sum_{i=1}^{t} x_{human,i}^{2} - \left(\sum_{i=1}^{t} x_{human,i}\right)^2}}$$

Where $x_{model,i}$ and $x_{human,i}$ refer to the dependent measure (e.g., Pmax, Pmax2) in the model and human data for trial i. The dependent measure has been averaged over all participants and problems for each trial. The t is the total number of trials in each of the model and human data sets. A value of $r = 1.0$ means that the model is able to completely explain the trend in the dependent variable of the human data, and a value of $r = 0.0$ means that the model is not able to explain the trend. A negative value of r means that the trend in human data is opposite to that in the model data.

Definition of Akaike Information Criterion (AIC)

$$\text{AIC} = t * \ln\frac{SSE}{t} + 2 * k$$

$$SSE = \sum_{i=1}^{t}(x_{model,i} - x_{human,i})^2$$

Where $x_{model,i}$ and $x_{human,i}$ refer to the average dependent measure (e.g., average Pmax2) in the model and human data over t trials of a task. The average in the dependent measure has been taken over all problems and participants. The SSE is the sum of squared errors between the human and model data set that is calculated for the average dependent measure. The t is the number of trials in the task, and k is the number of parameters in the model. The AIC measure incorporates both the effect of an MSD and the number of parameters in a model. The smaller or more negative the value of AIC, the better the respective model is. In the sampling paradigm where there is a single final consequential decision, the value of t equals 1 if the dependent measure is Pmax.

The Choice of Model Calibration Procedures

There are multiple procedures for finding optimal values for model parameters to enable calibration to human data. Two common mathematical methods classically used include the grid search and steepest-descent search (Busemeyer & Diederich, 2009). The basic idea in a grid search is to take a range of values of parameters with certain increments and then calculate the MSD between the model's predictions and human data on a dependent measure for all possible values of the parameters. The different parameter combinations results in a grid where an intersection in the grid is a set of parameter values to be run in a model. With just a few parameters, the grid search process might become time consuming and computationally complex. The steepest-descent search works by first randomly assuming values of the parameters (represented as a point in parameter space) and then calculating the MSD for all possible movements in different directions from the initial point. The next point to move to is the one that produced the largest decrease in the MSD. With more and more parameters in a model, the possible directions of movement might become very large and almost impossible to evaluate.

A more recent approach that was used in this article is to apply a Genetic Algorithm (GA; Holland, 1975) to find optimal parameters. The GA is different and better than the methods of optimization described above in several ways. The most important difference is that a GA works on a finite and limited "population of possible parameter values," whereas other methods only incorporate a single parameter in different iterations for evaluating the MSD. Another difference is that GA is probabilistic (stochastic) and not a deterministic process; thus, it has good chances of avoiding local optimal points in the parameter space (Jakobsen, 2010).

In our calibration process, the GA tries out different combinations of parameters to minimize the MSD between the model's proportion of maximization (Pmax or Pmax2) and the corresponding human's proportions across the six choice (SC) problems. Different parameter combinations (N) are selected and run in a model for a generation. Within a generation, a combination of parameters is used, and the MSD value is determined. The parameter combinations are then ranked from lowest (best) to highest (worst) based upon the calculated MSDs. After ranking, parameter combinations from the top half ranks are kept (N/2), and others (N/2) are discarded. The parameter combinations that are kept then duplicate themselves, bringing the number of parameter combinations back to the original amount (N). The N parameter combinations are then paired off with each other at random (thus forming N/2 pairs). Now, each parameter combination exchanges some of its adjustable parameter values with the corresponding parameter

(*Appendices continue*)

value of its partner (this is called "reproduction"). For example, suppose the following two three-parameter combinations have been paired off: (a1, b1, c1) and (a2, b2, c2). Due to the exchange of the adjustable parameter values a1 and a2 in the pair, the resulting parameter combinations will be (a2, b1, c1) and (a1, b2, c2). After the exchange, a new generation will start with a new set of *N* parameter combinations being run, and the process described above repeats. The stopping rule for the GA in different model optimizations was set at 10,000 generations. This value is extremely large and thus ensures a very high level of confidence in the values obtained.

Calibrating the IBL Model

The IBL model used in this article has three free parameters: decay d, noise σ, and probability of inertia *pInertia*. To calibrate the IBL model in the SC problems and the Technion Prediction Tournament's (TPT's) estimation set for both paradigms, we varied the d and σ parameters between 0.0 and 20.0 and the *pInertia* parameter between 0.0 and 1.0. This range of variation is large and ensures a high confidence in the resulting optimal parameters. The IBL model was run using the same number of simulated participants as the number of human participants that participated in the two paradigms for the SC problems and the TPT's estimation set (SC problems: 25 model participants per problem in the sampling paradigm and 24 model participants per problem over 400 trials in the repeated-choice paradigm; TPT's estimation set: 20 model participants per problem in the sampling paradigm and 20 model participants per problem over 100 trials in the repeated-choice paradigm).

Calibrating the ACT–R Model

The TPT's ACT–R model uses three parameters: decay d, noise σ, and *RT*. The d and σ parameters in this model have the same meaning as in the IBL model; in addition, the retrieval of instances from memory in this model is inhibited if the activation is below the *RT* parameter. If the activation of the most active instance that participates in blending is less than the *RT*, then the model is unable to retrieve the corresponding blended instance, and it uses a random decision to pick one of the two options. To calibrate the ACT–R model in the repeated-choice paradigm for the SC choice problems, we varied the d, σ parameters between 0.0 and 20.0, and the *RT* parameter between –20 and +20. This range of variation is large and ensures high confidence in the optimal parameters found. The model was run using the same number of simulated participants as the number of human participants that participated in the repeated-choice paradigm (i.e., 24 model participants per problem over 400 trials).

Calibrating the Ensemble Model

The TPT's Ensemble model uses 40 different parameters. The model does not explicitly sample the two options before making a final decision; rather, the decision-weight parameters assumed in the model allow it to incorporate the effect of sampling mathematically. Because finding the optimal values of 40 parameters becomes extremely complex and time consuming, we calibrated the 11 main parameters whose optimal values have also been listed in Erev et al. (2010) for the TPT's sampling paradigm. Six of the 11 parameters were varied between 0.0 and 10.0, and these included *Alpha*, *Beta*, *Gamma*, *Delta*, *Lambda*, and *Mu*. The other five parameters were varied between 0.0 and 1.0, as they represented decision-weights between 0.0 and 1.0. These five parameters included *Wmin*, *Wp*, *Sigma*, *Throg*, and *Thrp*. The Ensemble model was run using the same number of simulated participants as human participants that participated in the sampling paradigm (i.e., 25 model participants per problem).

(*Appendices continue*)

Appendix B

One Hundred Twenty Problems in the Technion Prediction Tournament's Estimation and Competition Data Sets

Table B1

The 60 Problems in the Estimation Set With the Average Proportion of Alternations (A-rate) and the Average Proportion of Risky Choices (R-rate) in the Repeated-Choice and Sampling Paradigms

			Risky		Safe		Repeated-choice		Sampling	
Problem	Domain	Probability	H	pH	L	M	A-rate	R-rate	A-rate	R-rate
1	Negative	High	−0.3	0.96	−2.1	−0.3	0.28	0.33	0.39	0.25
2	Negative	High	−0.9	0.95	−4.2	−1	0.1	0.5	0.41	0.55
3	Negative	Medium	−6.3	0.3	−15.2	−12.2	0.14	0.24	0.25	0.50
4	Negative	Medium	−10	0.2	−29.2	−25.6	0.13	0.32	0.34	0.30
5	Negative	Medium	−1.7	0.9	−3.9	−1.9	0.19	0.45	0.37	0.80
6	Negative	High	−6.3	0.99	−15.7	−6.4	0.18	0.68	0.36	0.75
7	Negative	Medium	−5.6	0.7	−20.2	−11.7	0.27	0.37	0.38	0.60
8	Negative	Low	−0.7	0.1	−6.5	−6	0.13	0.27	0.28	0.20
9	Negative	High	−5.7	0.95	−16.3	−6.1	0.12	0.43	0.35	0.60
10	Negative	Medium	−1.5	0.92	−6.4	−1.8	0.15	0.44	0.29	0.90
11	Negative	Low	−1.2	0.02	−12.3	−12.1	0.11	0.26	0.48	0.15
12	Negative	High	−5.4	0.94	−16.8	−6.4	0.17	0.55	0.44	0.65
13	Negative	Low	−2	0.05	−10.4	−9.4	0.1	0.11	0.34	0.20
14	Negative	Medium	−8.8	0.6	−19.5	−15.5	0.16	0.66	0.35	0.80
15	Negative	Low	−8.9	0.08	−26.3	−25.4	0.13	0.19	0.38	0.30
16	Negative	Low	−7.1	0.07	−19.6	−18.7	0.12	0.34	0.33	0.25
17	Negative	Low	−9.7	0.1	−24.7	−23.8	0.19	0.37	0.39	0.55
18	Negative	Medium	−4	0.2	−9.3	−8.1	0.22	0.34	0.39	0.40
19	Negative	Medium	−6.5	0.9	−17.5	−8.4	0.18	0.49	0.26	0.80
20	Negative	Medium	−4.3	0.6	−16.1	−4.5	0.1	0.08	0.39	0.20
21	Mixed	Low	2	0.1	−5.7	−4.6	0.08	0.11	0.36	0.20
22	Mixed	Medium	9.6	0.91	−6.4	8.7	0.15	0.41	0.33	0.70
23	Mixed	Medium	7.3	0.8	−3.6	5.6	0.12	0.39	0.33	0.70
24	Mixed	Low	9.2	0.05	−9.5	−7.5	0.07	0.08	0.26	0.05
25	Mixed	Low	7.4	0.02	−6.6	−6.4	0.11	0.19	0.38	0.10
26	Mixed	Low	6.4	0.05	−5.3	−4.9	0.1	0.2	0.30	0.15
27	Mixed	High	1.6	0.93	−8.3	1.2	0.14	0.5	0.37	0.70
28	Mixed	Medium	5.9	0.8	−0.8	4.6	0.26	0.58	0.32	0.65
29	Mixed	High	7.9	0.92	−2.3	7	0.14	0.51	0.30	0.65
30	Mixed	Medium	3	0.91	−7.7	1.4	0.18	0.41	0.29	0.70
31	Mixed	High	6.7	0.95	−1.8	6.4	0.11	0.52	0.34	0.70
32	Mixed	High	6.7	0.93	−5	5.6	0.11	0.49	0.39	0.55
33	Mixed	High	7.3	0.96	−8.5	6.8	0.08	0.65	0.41	0.75
34	Mixed	Low	1.3	0.05	−4.3	−4.1	0.1	0.3	0.27	0.10
35	Mixed	High	3	0.93	−7.2	2.2	0.11	0.44	0.27	0.55
36	Mixed	Low	5	0.08	−9.1	−7.9	0.07	0.09	0.45	0.20
37	Mixed	Medium	2.1	0.8	−8.4	1.3	0.23	0.28	0.39	0.35
38	Mixed	Low	6.7	0.07	−6.2	−5.1	0.14	0.29	0.36	0.20
39	Mixed	Medium	7.4	0.3	−8.2	−6.9	0.2	0.58	0.35	0.70
40	Mixed	High	6	0.98	−1.3	5.9	0.15	0.61	0.20	0.70
41	Positive	Medium	18.8	0.8	7.6	15.5	0.11	0.52	0.42	0.60
42	Positive	High	17.9	0.92	7.2	17.1	0.07	0.48	0.24	0.80
43	Positive	Low	22.9	0.06	9.6	9.2	0.07	0.88	0.32	0.90
44	Positive	High	10	0.96	1.7	9.9	0.11	0.56	0.34	0.70
45	Positive	Medium	2.8	0.8	1	2.2	0.2	0.48	0.19	0.70
46	Positive	Low	17.1	0.1	6.9	8	0.12	0.32	0.44	0.20
47	Positive	Low	24.3	0.04	9.7	10.6	0.13	0.25	0.31	0.20
48	Positive	High	18.2	0.98	6.9	18.1	0.14	0.59	0.33	0.75
49	Positive	Medium	13.4	0.5	3.8	9.9	0.13	0.13	0.35	0.45
50	Positive	Low	5.8	0.04	2.7	2.8	0.16	0.35	0.35	0.20
51	Positive	High	13.1	0.94	3.8	12.8	0.09	0.52	0.41	0.65

(Appendices continue)

Table B1 (*continued*)

Problem	Domain	Probability	Risky H	Risky pH	Safe L	Safe M	Repeated-choice A-rate	Repeated-choice R-rate	Sampling A-rate	Sampling R-rate
52	Positive	Low	3.5	0.09	0.1	0.5	0.12	0.26	0.29	0.25
53	Positive	Low	25.7	0.1	8.1	11.5	0.09	0.11	0.35	0.25
54	Positive	Low	16.5	0.01	6.9	7	0.14	0.18	0.30	0.25
55	Positive	High	11.4	0.97	1.9	11	0.1	0.66	0.33	0.70
56	Positive	High	26.5	0.94	8.3	25.2	0.1	0.53	0.41	0.50
57	Positive	Medium	11.5	0.6	3.7	7.9	0.27	0.45	0.30	0.45
58	Positive	High	20.8	0.99	8.9	20.7	0.17	0.63	0.31	0.65
59	Positive	Medium	10.1	0.3	4.2	6	0.19	0.32	0.34	0.45
60	Positive	High	8	0.92	0.8	7.7	0.14	0.44	0.38	0.55
Average							0.14	0.40	0.34	0.49

Note. H = high outcome; pH = high outcome with some probability; L = low outcome; M = medium outcome.

Table B2

The 60 Problems in the Competition Set With the Average Proportion of Alternations (A-rate) and the Average Proportion of Risky Choices (R-rate) in the Repeated-Choice and Sampling Paradigms

Problem	Domain	Probability	Risky H	Risky pH	Safe L	Safe M	Repeated-choice A-rate	Repeated-choice R-rate	Sampling A-rate	Sampling R-rate
1	Negative	Low	−8.7	0.06	−22.8	−21.4	0.17	0.25	0.35	0.45
2	Negative	Low	−2.2	0.09	−9.6	−8.7	0.17	0.27	0.31	0.15
3	Negative	Low	−2	0.1	−11.2	−9.5	0.16	0.25	0.34	0.10
4	Negative	Low	−1.4	0.02	−9.1	−9	0.14	0.33	0.38	0.20
5	Negative	Low	−0.9	0.07	−4.8	−4.7	0.13	0.37	0.27	0.35
6	Negative	High	−4.7	0.91	−18.1	−6.8	0.21	0.63	0.26	0.75
7	Negative	Low	−9.7	0.06	−24.8	−24.2	0.17	0.30	0.33	0.50
8	Negative	High	−5.7	0.96	−20.6	−6.4	0.17	0.66	0.33	0.65
9	Negative	Low	−5.6	0.1	−19.4	−18.1	0.09	0.31	0.29	0.20
10	Negative	Medium	−2.5	0.6	−5.5	−3.6	0.12	0.34	0.36	0.50
11	Negative	High	−5.8	0.97	−16.4	−6.6	0.12	0.61	0.18	0.65
12	Negative	Low	−7.2	0.05	−16.1	−15.6	0.09	0.25	0.20	0.40
13	Negative	High	−1.8	0.93	−6.7	−2	0.11	0.44	0.25	0.55
14	Negative	Medium	−6.4	0.2	−22.4	−18	0.15	0.20	0.30	0.15
15	Negative	High	−3.3	0.97	−10.5	−3.2	0.1	0.16	0.27	0.10
16	Negative	Medium	−9.5	0.1	−24.5	−23.5	0.12	0.39	0.41	0.70
17	Negative	High	−2.2	0.92	−11.5	−3.4	0.13	0.47	0.27	0.65
18	Negative	High	−1.4	0.93	−4.7	−1.7	0.09	0.41	0.22	0.55
19	Negative	Medium	−8.6	0.1	−26.5	−26.3	0.18	0.49	0.27	0.60
20	Negative	Low	−6.9	0.06	−20.5	−20.3	0.14	0.25	0.21	0.60
21	Mixed	Medium	1.8	0.6	−4.1	1.7	0.10	0.08	0.41	0.10
22	Mixed	High	9	0.97	−6.7	9.1	0.11	0.14	0.30	0.15
23	Mixed	Low	5.5	0.06	−3.4	−2.6	0.15	0.28	0.27	0.20
24	Mixed	High	1	0.93	−7.1	0.6	0.16	0.46	0.40	0.65
25	Mixed	Medium	3	0.2	−1.3	−0.1	0.13	0.21	0.38	0.25
26	Mixed	Medium	8.9	0.1	−1.4	−0.9	0.12	0.23	0.27	0.25
27	Mixed	High	9.4	0.95	−6.3	8.5	0.14	0.67	0.36	0.55
28	Mixed	High	3.3	0.91	−3.5	2.7	0.17	0.58	0.34	0.65
29	Mixed	Medium	5	0.4	−6.9	−3.8	0.17	0.39	0.24	0.70
30	Mixed	Low	2.1	0.06	−9.4	−8.4	0.12	0.33	0.28	0.30
31	Mixed	Medium	0.9	0.2	−5	−5.3	0.09	0.88	0.29	0.95
32	Mixed	Low	9.9	0.05	−8.7	−7.6	0.06	0.21	0.21	0.30
33	Mixed	Low	7.7	0.02	−3.1	−3	0.10	0.28	0.28	0.35
34	Mixed	High	2.5	0.96	−2	2.3	0.13	0.52	0.23	0.50
35	Mixed	High	9.2	0.91	−0.7	8.2	0.09	0.56	0.26	0.60

(*Appendices continue*)

Table B2 (*continued*)

Problem	Domain	Probability	Risky		Safe		Repeated-choice		Sampling	
			H	pH	L	M	A-rate	R-rate	A-rate	R-rate
36	Mixed	High	2.9	0.98	−9.4	2.9	0.23	0.34	0.27	0.35
37	Mixed	Low	2.9	0.05	−6.5	−5.7	0.17	0.30	0.19	0.35
38	Mixed	High	7.8	0.99	−9.3	7.6	0.09	0.62	0.26	0.75
39	Mixed	Medium	6.5	0.8	−4.8	6.2	0.08	0.32	0.25	0.35
40	Mixed	High	5	0.9	−3.8	4.1	0.08	0.46	0.23	0.50
41	Positive	High	20.1	0.95	6.5	19.6	0.20	0.50	0.28	0.65
42	Positive	Medium	5.2	0.5	1.4	5.1	0.10	0.08	0.32	0.05
43	Positive	Medium	12	0.5	2.4	9	0.16	0.17	0.37	0.25
44	Positive	High	20.7	0.9	9.1	19.8	0.19	0.44	0.34	0.55
45	Positive	Low	8.4	0.07	1.2	1.6	0.12	0.20	0.31	0.25
46	Positive	Medium	22.6	0.4	7.2	12.4	0.20	0.41	0.27	0.30
47	Positive	High	23.4	0.93	7.6	22.1	0.14	0.72	0.40	0.65
48	Positive	Low	17.2	0.09	5	5.9	0.12	0.24	0.42	0.50
49	Positive	Medium	18.9	0.9	6.7	17.7	0.08	0.57	0.47	0.45
50	Positive	Low	12.8	0.04	4.7	4.9	0.06	0.26	0.25	0.30
51	Positive	Low	19.1	0.03	4.8	5.2	0.07	0.22	0.18	0.25
52	Positive	High	12.3	0.91	1.3	12.1	0.12	0.41	0.36	0.35
53	Positive	Medium	6.8	0.9	3	6.7	0.11	0.41	0.28	0.40
54	Positive	Medium	22.6	0.3	9.2	11	0.15	0.60	0.33	0.85
55	Positive	Low	6.4	0.09	0.5	1.5	0.12	0.28	0.31	0.40
56	Positive	Low	15.3	0.06	5.9	7.1	0.10	0.17	0.18	0.25
57	Positive	Medium	5.3	0.9	1.5	4.7	0.12	0.66	0.29	0.65
58	Positive	Medium	21.9	0.5	8.1	12.6	0.10	0.47	0.30	0.80
59	Positive	Medium	27.5	0.7	9.2	21.9	0.11	0.42	0.21	0.25
60	Positive	Medium	4.4	0.2	0.7	1.1	0.12	0.38	0.27	0.70
Average							0.13	0.38	0.29	0.44

Note. H = high outcome; pH = high outcome with some probability; L = low outcome; M = medium outcome.

Received February 24, 2011
Revision received May 16, 2011
Accepted May 27, 2011 ■