Original Research

Accounting for outcome and process measures in dynamic decision-making tasks through model calibration

Varun Dutt[1] and Cleotilde Gonzalez[2]

[1]School of Computing and Electrical Engineering and School of Humanities and Social Sciences, Indian Institute of Technology Mandi, India and [2]Dynamic Decision Making Laboratory, Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA, USA

Computational models of learning and the theories they represent are often validated by calibrating them to human data on decision outcomes. However, only a few models explain the process by which these decision outcomes are reached. We argue that models of learning should be able to reflect the process through which the decision outcomes are reached, and validating a model on the process is likely to help simultaneously explain both the process as well as the decision outcome. To demonstrate the proposed validation, we use a large dataset from the Technion Prediction Tournament and an existing Instance-based Learning Model. We present two ways of calibrating the Model's parameters to human data: on an outcome measure and on a process measure. In agreement with our expectations, we find that calibrating the Model on the process measure helps to explain both the process and outcome measures compared to calibrating the Model on the outcome measure. These results hold when the Model is generalized to a different dataset. We discuss implications for explaining the process and the decision outcomes in computational models of learning.

Keywords: outcome and process measures, computational models of learning, Instance-based learning, dynamic decisions, binary choice, calibration

Unlike disciplines like economics, models of decision making in psychology often incorporate theories of the underlying cognitive processes that lead to specific outcomes in a decision task. For example, Instance-Based Learning Theory (IBLT; Gonzalez & Dutt, 2011), a theory of how people make dynamic decisions commonly includes assumptions of how people search for information (i.e., the process) and how this information search helps people to arrive at a decision (i.e., the outcome). However, many of the process theories and corresponding models are just tested on an outcome level; rather than on the process level itself (Johnson et al., 2008). Accounting for both the decision outcomes and the process through which these outcomes are reached is important in mathematical models (Scheres & Sanfey, 2006). That is because by accounting for the process and decision outcomes will enable such models to provide better account of the observed phenomena. Furthermore, it is also important to account for process and decision outcomes in computational models of learning that try to explain human decisions (Busemeyer & Diederich, 2009, Erev & Barron, 2005, Rapoport

& Budescu, 1992). For example, researchers investigating choice behavior are often interested in explaining the overall maximization behavior (an outcome measure) and the exploratory behavior (e.g., alternation between alternatives, a process measure) through cognitive models, which explains how people learn to maximize long-term rewards (Biele, Erev & Ert, 2009; Erev, Ert, Roth, Haruvy et al., 2010; Gonzalez & Dutt, 2011).

Amidst the importance of accounting for both the decision outcome and the process, literature has revealed a strong relationship between these two, where the resulting outcome is consistent with the adopted process (Erev & Barron, 2005; Green, Price & Hamburger, 1995; Hills & Hertwig, 2010). According to Erev and Barron (2005), one expects a strong relationship between process and decision outcomes in cases where the decision environment is dynamic (i.e., repeated), and where the decision outcome is contingent upon the process. For example, consider a repeated binary-choice task, where choices are made repeatedly between two alternatives. One of the alternatives is risky with a high outcome and a low outcome. These two outcomes occur with a certain pre-defined probabilities when this risky alternative is chosen. The other alternative is safe with a medium outcome. This medium outcome occurs with a sure (100%) chance when this alternative is chosen. Now, if the expected value of the risky alternative is greater than that of the safe alternative (i.e., the safe alternative is maximizing), then participants who alternate a lot while selecting alternatives would end-up maximizing their choices only half of the time. In fact, Hills and Hertwig (2010) show that people seem to rely on two distinct alternation processes while making binary choices; both these processes achieve different amounts of maximization behavior. These arguments are not only relevant to human decisions but also to decision making in animals. For example, Green et al. (1995) have shown that pigeons can only learn to maximize their outcomes by alternating between available alternatives in a probabilistic environment involving repeated choices between safe and risky alternatives.

Calibrating models to both process and outcome measures from one-time sequential sampling tasks is already common in literature (Ratcliff, 1978; Ratcliff & Smith, 2004). For example, Ratcliff (1978) calibrated models to

---

Corresponding author: Varun Dutt, School of Computing and Electrical Engineering and School of Humanities and Social Sciences, Indian Institute of Technology, Mandi, PWD Rest House, Near Bus Stand, Mandi – 175 001, Himachal Pradesh, India. e-mail: varun@iitmandi.ac.in

both outcome and process measures in an old-new recognition memory task. In this task, the outcome measure was proportion of correct responses and the process measure was the accumulation of evidence to a threshold for making a response. In fact, calibrating models to both outcome and process measures in one-time choice tasks is so commonthat a suite of software called Diffusion Model Analysis Toolbox (DMAT, Vandekerckhove & Tuerlinckx, 2007) has been recently developed for this purpose.

In contrast, to the authors' best knowledge, except for one study (mentioned below) no one has explicitly calibrated models to outcome and process measures simultaneously in dynamic decision making tasks (Johnson, Schulte-Mecklenbeck & Willemsen, 2008). Johnson et al. (2008) demonstrated via computational modeling that the priority heuristic, which provides a novel account of how people make risky choices, captures the decision outcomes; yet, this heuristic fails to account for the process measures. The general finding is that although certain behavioral results reveal a strong connection between the decision outcome and the process, existing models of learning in dynamic decision tasks rarely show any relationship between them (Dember & Fowler, 1958; Erev & Barron, 2005; Erev, Ert, Roth, Haruvy et al., 2010; Rapoport & Budescu, 1992; Rapoport, Erev, Abraham & Olson, 1997; Tolman, 1925). For example, although the outcome results (i.e., maximization) in a symmetrical zero-sum matching pennies game were consistent with predictions from a reinforcement-learning algorithm, process results (i.e., alternations between alternatives) could not be accounted for by the algorithm (Erev & Barron, 2005; Rapoport & Budescu, 1992). Similarly, according to Johnson et al. (2008), the priority heuristic, a strategy to account for risky choices, fails to account for the process measures in dynamic decision tasks.

In one study, Gonzalez and Dutt (2011) have calibrated cognitive models in the sampling paradigm (a dynamic task), where participants are asked to sample options free of cost before making a consequential choice for real. Gonzalez and Dutt (2011) demonstrate that a computational model based upon the IBLT (Gonzalez, Lerch & Lebiere, 2003), ("IBL model" hereafter), when calibrated on the outcome measure, was able to also explain the process measure better than the best models known in two different experimental paradigms. Gonzalez and Dutt (2011) however, did not calibratetheir model on the process measure as well. Thus, it remains unclear what effect calibrating a model to the process measure compared to the outcome measure has on the model predictions of both these measures. In general, one expects the decision outcome to be the result of the process (Johnson et al., 2008). Thus, calibrating models on process measures rather than outcome measures should have benefits in explaining both these measures at the same time.

Although it is hard to find models calibrated to outcome and process measures in dynamic tasks, past studies have made certain qualitative predictions of dynamic decision models (Busemeyer, 1985; Hertwig, Barron, Weber & Erev, 2004; Lee, Zhang, Munro & Steyvers, 2011) on outcome and process measures. However, a quantitative empirical investigation of these models on both these measures is something currently lacking and much needed in literature. This paper makes a contribution in this area by investigating the benefit of calibrating cognitive models to outcome and process data in a dynamic decision task.

In this paper, we evaluate the role of calibrating a computational model to either the decision outcome or the process in explaining and predicting both these measures. Specifically, we calibrate an IBL model (Gonzalez & Dutt, 2011), to a risk-taking measure (decision outcome) or an alternation measure (process), and evaluate the model fits to human data (through parameter calibration in a dataset) and predictions (through generalization in a dataset different from calibration). Given the hypothesized benefits of calibrating models on process measures (Camerer & Ho, 1999; Suppes & Atkinson 1959), we expect that the IBL model being calibrated to the alternation measure would improve its explanation about both the risk-taking and alternations compared to when it is calibrated on the risk-taking measure. We use two large human datasets, estimation and competition, that were collected for the 2008 Technion Prediction Tournament (TPT (Erev, Ert, Roth, Haruvy et al., 2010). The choice of TPT datasets is because the main focus of the tournament was on outcome measures, where no attention was given to process measures (Erev, Ert, Roth, Haruvy et al., 2010). That is because it was felt that paying less attention to the process measures can actually help the prediction of the outcome measures (Erev & Haruvy, 2005; Estes, 1962), which is contrary to the hypothesis under test in this paper. Thus, this dataset becomes an ideal choice for testing a process-measure calibrated model's ability to perform on the outcome measure. In what follows, we first discuss the role of the calibration process in computational models. Next, we present the effects of calibrating an existing IBL model on the outcome measure or the process measure on the explanations and predictions of one or both measures in the TPT's datasets. We close this paper by discussing the role of model calibration to account for both the process and decision outcomes.

The Role of Model Calibration in Explaining Different Measures of Performance Calibrating a model to human data means finding the values of its parameters that minimize the deviation between model's predictions and observations on a dependent measure. In the TPT, several influential models[1] of learning in binary choice were calibrated and evaluated on only the outcome measure (risk-taking) and not on the process measure (alternations). These models were able to account for risk-taking very well; however, many of them did not provide any way of computing the alternations (Gonzalez & Dutt, 2011). In fact, most of the competing models did not provide any way to explain the learning process (see an extended discussion about these models in Gonzalez and Dutt (2011)). For example, a number of models submitted to the TPT used prospect theory (Tversky & Kahneman, 1992), to predict choices based upon calibrated mathematical functions. Prospect Theory does not provide any mechanism that would predict the sequential selection of options over time. In fact, only a few recent models of repeated binary-choice may account for both the risk-taking and alternation measures simultaneously: One of these models is the Inertia Sampling and Weighting (I-SAW) model (Chen et al., 2011; Nevo & Erev, 2012; Erev, Ert, Roth, Haruvy et al., 2010) and the other is an IBL model (Gonzalez & Dutt, 2011; Gonzalez, Dutt & Lejarraga, 2011; Lejarraga, Dutt & Gonzalez, 2012). However, these models were calibrated on both the outcome and process measures at the same time, which makes it

---

[1]Some of these models included the two-stage sampler model, the normalized reinforcement learning with inertia model, and the explorative sampler with recency model (Erev, Ert, Roth, Haruvy et al., 2010)

difficult to evaluate the utility of calibrating models to one of these measures.

We expect that calibrating a model to the process measure should generally be beneficial for the model's ability to explain both the process and outcome measures upon generalization to novel conditions. Next, we provide details about the TPT datasets that we use to evaluate the IBL model.

Method

Risk-taking and Alternations in Technion Prediction Tournament

Competing models submitted to the TPT were evaluated according to the generalization criterion method (Busemeyer & Wang, 2000), by which models were calibrated on choices made by participants in 60 problems (the estimation set) and later tested in a new set of 60 problems (the competition set) with the parameters obtained from the calibration process in the estimation set. The generalization criterion method was believed to be a true test for models to explain observed choice decisions. Although the TPT involved three different experimental paradigms, we only use data from the "E-repeated" paradigm that involved consequential choices in a repeated binary-choice task with immediate outcome feedback on the chosen alternative. For each of the 60 problems in the estimation and competition sets in this paradigm, a sample of 100 participants was randomly assigned into 5 groups of 20 participants each, and each group completed 12 of the 60 problems. Each participant was instructed to repeatedly and consequentially select between two unlabeled buttons on a computer screen in order to maximize long-term rewards for a block of 100 trials per problem (this end point was not known to participants). One button was associated with a risky alternative and the other button with a safe alternative. Selecting an alternative, safe or risky, generated an outcome for the selected alternative (thus, the foregone outcome on the unselected alternative was not shown). The selection of the alternative with the higher expected value, which could be either the safe or risky button, would maximize a participant's long-term rewards. Therefore, choosing a maximizing alternative across all the repeated trials would constitute the optimal strategy in the task. Other details about the E-repeated paradigm are reported in Erev, Ert, Roth, Haruvy et al. (2010).

The models submitted to the TPT were not provided with human data for alternation between options (i.e., the A-rate or the process measure), and they were evaluated only according to their ability to account for the risk-taking behavior (i.e., the R-rate or the outcome measure) (Erev, Ert, Roth, Haruvy et al., 2010). We calculated the A-rate for analyses of alternations from the TPT data (see results in Gonzalez and Dutt, 2011). First, alternations are either coded as 1s, the respondent switched from making a risky or safe choice in the last trial to making a safe or risky choice in the

current trial; or they are coded as 0s, the respondent simply repeated the last trial's choice. The proportion of alternations in each trial is computed by averaging the alternations over 20 participants per problem and the 60 problems in each dataset. The R-rate is the proportion of risky choices in each trial averaged over 20 participants per problem and the 60 problems in each dataset. A problem is defined as consisting of two alternatives, risky and safe. In the risky alternative, there are two possible outcomes, high and low, where the occurrence of these outcomes is determined by corresponding probability value. In the safe alternative, there is one possible outcome, medium, where this outcomes occurs with a 100% chance. For calculating the A-rate and R-rate, the averaging is done over 20 participants as this many participants were collected in the TPT (Erev, Ert, Roth, Haruvy et al., 2010).

Figure 1 shows the overall R-rate and A-rate over 99 trials from trial 2 to trial 100 in the estimation and competition sets. As seen in both of these datasets, the R-rate is relatively constant across trials, in contrast to the sharp decrease in the A-rate. The sharp decrease in the A-rate shows a transition in the pattern of information-search across trials (Gonzalez & Dutt, 2011). Overall, these R-rate and A-rate curves suggest that risk-taking remains relatively steady across trials, while they learn to alternate less and choose one of the two alternatives more often. Thus, the A-rate (process) is more dynamic compared to the R-rate (decision outcome) and due to these differences it is likely to be harder for a model to account for the A-rate compared to the R-rate. We use the R-rate and A-rate curves in Figure 1 to evaluate the role of model calibration ahead in this paper.

An Instance-based Learning Model of Repeated Binary-choice

IBLT (Gonzalez et al., 2003) has been used as the basis for developing computational models that capture human behavior in a wide variety of dynamic decision making tasks. These include dynamically-complex tasks like the water purification plant task (Gonzalez & Lebiere, 2005; Gonzalez et al., 2003; Martin, Gonzalez & Lebiere, 2004), training paradigms of simple and complextasks (Gonzalez, Best, Healy, Bourne & Kole, 2010), simple stimulus-response practice and skill acquisition tasks (Dutt, Yamaguchi, Gonzalez & Proctor, 2009) and repeated binary-choice tasks (Gonzalez & Dutt, 2011; Gonzalez et al., 2011; Lebiere, Gonzalez & Martin, 2007; Lejarraga et al., 2012) among others. The different computational applications of IBLT illustrate its generality and ability to capture decisions from experience in multiple contexts.

A recent IBL model has showcased the theory's robustness across multiple choice tasks: A probability-learning task, a repeated binary-choice task with fixed probabilities, and a repeated binary-choice task with changing probabilities (Lejarraga et al., 2012). We use this model to evaluate the effects of model cali-
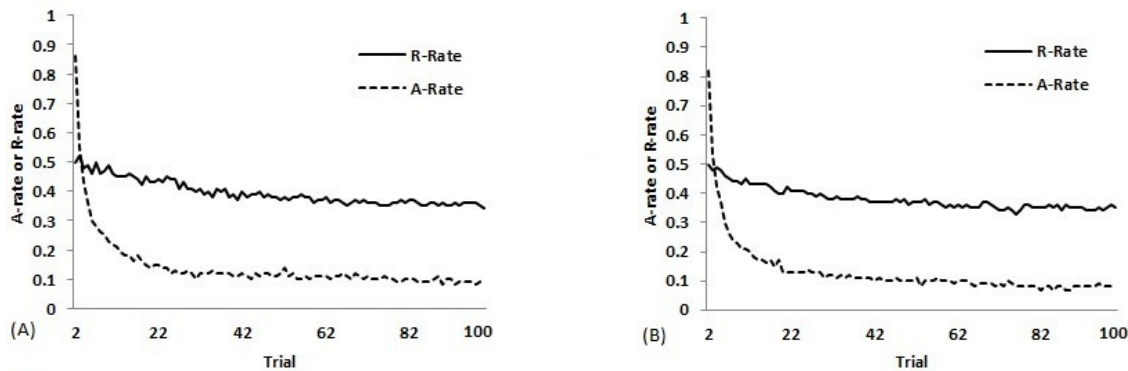
Figure 1. (A) The R-rate and A-rate across trials observed in human data in the estimation set of the TPT between trial 2 and trial 100. (B) The R-rate and A-rate across trials observed in human data in the competition set of the TPT between trial 2 and trial 100.

bration to different outcome or process measures. The model's formulations and decision-making process are further explained in other publications (Gonzalez & Dutt, 2011; Lejarraga et al., 2012) and summarized in the Appendix. This model makes choice selections between alternatives in a trial by comparing the weighted averages of observed outcomes on each alternative called "blended values." A blended value for an alternative, safe or risky, is a function of the probability of retrieving instances from memory multiplied by their respective outcomes that have been observed on previous selections of the alternative (Lebiere, 1999; Lejarraga et al., 2012). Each instance consists of a label that identifies a decision alternative in the task and the outcome obtained. For example, (risky, $32) is an instance where the decision was to choose the risky alternative and the outcome obtained was $32. The probability of retrieving an instance from memory, which is used to compute the blended value, is a function of its activation (Anderson & Lebiere, 1998). Each observed outcome (represented by a corresponding instance in memory) has an activation value that is a function of the recency and frequency of observing the outcome plus a noise term. This simplified activation equation has shown to be sufficient at explaining human choices in several experiential tasks (Gonzalez & Dutt, 2011; Lejarraga et al., 2012). The activation is influenced by the decay parameter , which captures the rate of forgetting or the reliance on recency and frequency of observing outcomes. The higher the value of the parameter, the greater is the model's reliance on outcomes experienced recently. The activation is also influenced by a noise parameter that is important for capturing the variability in human behavior from one participant to another. IBL borrows and s parameters and the activation equation from a popular cognitive framework called ACT-R (Atomic Components of Thought – Rational; Anderson & Lebiere, 1998). However, unlike ACT-R where and s parameters are kept fixed, we calibrate the values of these parameters in the IBL model to account for choices in human data. The model equations for blending and activation are included in the Appendix.

Results

Model Calibration to Different Measures

We used a genetic algorithm program to calibrate the model's parameters to minimize the mean squared deviation (MSD) between its predictions and the observed average A-rate per problem or the average R-rate per problem. The average R-rate per problem and the average A-rate per problem were computed by averaging the risky choices and alternations in each problem over 20 participants per problem and 100 trials per problem (for a problem's definition, please see the description above). Later, the MSDs were calculated across the 60 estimation set problems by using the average R-rate per problem and by the average A-rate per problem from the model and human data. For calibration, both the s and the d parameters were varied between 0.0 and 10.0 and the genetic algorithm was run for 500 generations (crossover rate = 50%; mutation rate = 10%). The assumed range of variation for the s and d parameters and the number of generations in the genetic algorithm is large, and it ensures that the optimization process does not miss the minimum MSD value due to a small range of parameter variation (for more details about genetic algorithm optimization, please see Gonzalez & Dutt, 2011). We calibrated the IBL model separately on the R-rate and the A-rate measures, and the optimized values of the d and s parameters were determined for each calibration.

The model calibrated on the R-rate produced the smallest MSD for $d = 5.00$ and $s = 1.50$. These parameters have the same optimal values as reported by Lejarraga et al. (2012), who had also calibrated this IBL model on the R-rate measure on the same dataset. As documented by Lejarraga et al. (2012), the value of both the d and s parameters is high compared to the ACT-R default values of $d = 0.5$ and $s = 0.25$ (Anderson & Lebiere, 1998). Furthermore, the model calibrated on the A-rate produced the smallest MSD for $d = 9.74$ and $s = 0.96$. Thus, calibrating the model on the A-rate produces a greater value for the d parameter and a slightly smaller value for the s parameter. The greater d parameter value suggests a high

dependency on recently experienced outcomes to make choice decisions.

Figure 2 shows the MSDs for the R-rate and the A-rate from the IBL model that was calibrated on the R-rate or the A-rate in the estimation set. When model parameters were calibrated on the R-rate (i.e., $d = 5.0$ and $s = 1.5$), the model explained the R-rate quite well (MSD = 0.008), but it explained the A-rate less well (MSD = 0.063). Thus, the model explains the outcome measure well when calibrated on the outcome measure; but, it explains the process measure less well. In contrast, when the IBL model parameters are calibrated on the A-rate, the model explains the A-rate much better (MSD=0.002) and the resulting R-rate also relatively well (MSD = 0.023). Thus, the benefit of calibrating the model on the A-rate measure (= 0.061) is larger than the detriment of calibrating the model on the R-rate measure (= 0.015). Overall, these results show that by calibrating the IBL model to the process measure, one is able to explain both the process and outcome measures better than by calibrating the IBL model to the outcome measure. Thus, these results suggest that the components of the IBL model are good representations of the A-rate process and well as the R-rate decision outcomes, especially when accounting for the A-rate is more challenging than the R-rate because the A-rate is more dynamic than the R-rate (Gonzalez & Dutt, 2011).
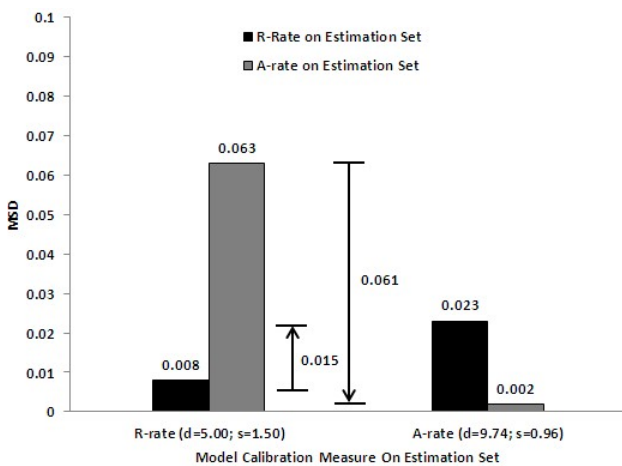
Figure 2. The MSD for the R-rate per problem and A-rate per problem in the estimation set of the TPT. The model was either calibrated on the R-rate per problem or calibrated on the A-rate per problem in the estimation set. The calibrated values of the d and s parameters obtained for each measure(R-rate or A-rate per problem) have been shown in brackets.The differences for calibrating with A-rate measure(respective R-rate measure) are shown by two vertical arrows.

Figure 3 presents the human and model R-rate and A-rate across trials when the model was calibrated to the R-rate (Figure 3A) and when it was calibrated to the A-rate (Figure 3B). Here, it can be observed how the model explains the human learning data better for the measure used to calibrate the model.

Generalizing the Calibrated IBL model to the Competition set

The demonstration that calibrating a model to a process measure helps explain both the process and outcome measures is an important way to corroborate the consistency of predictions from cognitive models. A robust model should be able to explain the learning process, as well as the outcomes resulting from that very process.

According to Lebiere, Gonzalez, and Warwick (2009), models that explain only the outcome and not the process behavior might find it difficult to generalize their predictions to novel conditions. Here, we used the generalization criterion test (Ahn, Busemeyer, Wagenmakers, & Stout, 2009; Busemeyer & Wang, 2000), to investigate the predictions that the different calibration procedures can make in novel data sets: We ran the calibrated models in novel conditions to evaluate and compare performance. The model calibrated to the TPT's estimation set on the R-rate or the A-rate was generalized to TPT's competition set by keeping the same parameter values that were derived during calibration. The model was run using 20 participants per problems and 60 problems in the competition set. There were different sets of problems used between the estimation and competition sets. Also, these problems were run as part of two separate experiments involving different human participants. Given these differences, one expects poorer performance from both the models in the competition set compared to the estimation set. However, as the algorithm used to generate problems in the competition set was same as that used to generate problems in the estimation set, one also expects both models to showcase results that are similar to those found for the estimation set: The model calibrated to the process measure is able to explain both the process and outcome measures better than the model calibrated to the outcome measure.

Figure 4 shows the resulting MSDs from generalizing the IBL model to the competition set. The model that was calibrated on the estimation set's R-rate resulted in the best predictions for the same measure in the competition set (MSD = 0.006); however, its predictions for the A-rate were relatively inferior (MSD = 0.074). Furthermore, the model that was calibrated on the A-rate resulted in the best predictions for the same measure in the competition set (MSD = 0.006) with reasonably good predictions for the R-rate (MSD = 0.032). Thus, again the improvement in MSD for the A-rate is larger than (= 0.068) the decrement in the MSD for the R-rate (= 0.026). Also note that the results in competition set (Figure 4) generate poorer performance (higher MSDs) from the models, in general, compared to those in the estimation set (Figure 2).

As in the estimation set, these results translate to the process of learning over trials (see Figure 5). The model's predictions for the measure on which it was calibrated to in the estimation set are the best. The model that was calibrated on the R-rate in the estimation set predicted the R-rate better than the A-rate
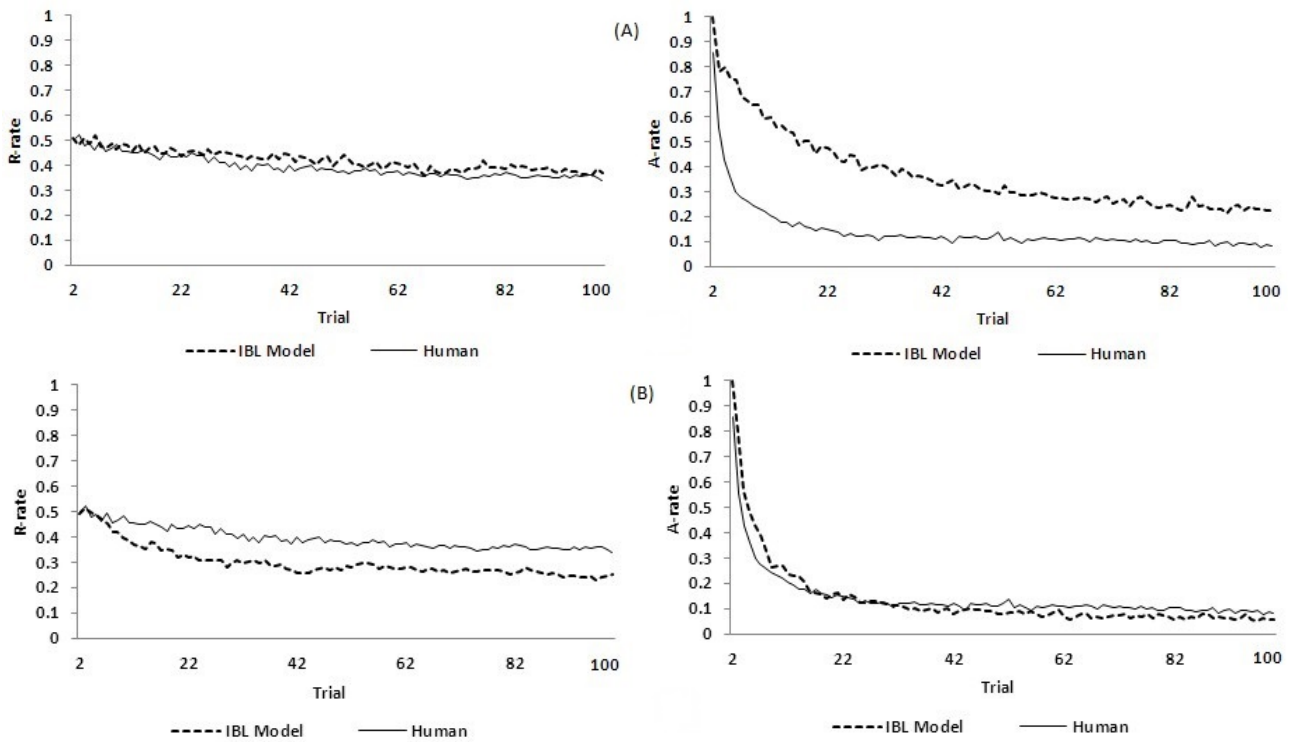
Figure 3. The R-rate and A-rate across trials predicted by the IBL model and that observed in human data in the TPT's estimation set. Panels A and B show the results of calibrating the IBL model to the R-rate per problem and the A-rate per problem, respectively.
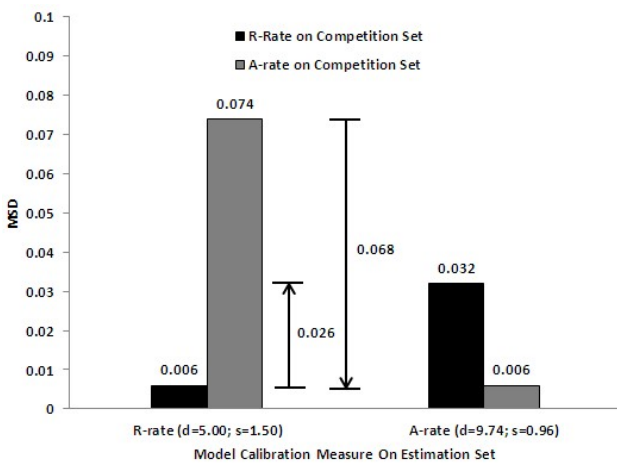

Figure 4. The MSD for the R-rate per problem and A-rate per problem in the competition set of the TPT. The model was either calibrated on the R-rate per problem or calibrated on the A-rate per problem in the estimation set. The calibrated values of the d and s parameters obtained for each measure (R-rate or A-rate per problem)in the estimation set have been shown in brackets. The differences for calibrating with A-rate measure (respective R-rate measure) are shown by two vertical arrows.

(Figure 5A); however, the model that was calibrated on the A-rate in the estimation set predicted both the R-rate and A-rate over time quite well (Figure 5B).

Discussion

We argue that strong and robust models of human behavior need to explain both the decision outcome and the process from which that outcome came about. We suggest that many models of human behavior, particularly in the context of repeated choice and dynamic decisions from experience, have only focused on predicting outcomes but not the process. Furthermore, most of the existing computational models of experiential decisions explain the decision outcomes, while completely ignoring or failing to account for the process through which these decision outcomes are reached (see a review of models in (Gonzalez & Dutt, 2011). This observation is perhaps not a coincidence, because predicting outcome as a result of a process is very challenging (Erev & Barron, 2005; Rapoport et al., 1997).

Our findings presented the robustness of explaining and predicting outcome and process measures through an IBL model. We demonstrated a method for finding out a cognitive model's ability in explaining both the process and the decision outcomes. The model's calibration on the process measure reduced the MSD for the A-rate (process) by a large amount without a large deterioration in the MSD for the R-rate (decision outcome). The proposed calibration was also helpful in accounting for both these measures after the model was generalized into a novel condition.

Explaining both the process and decision outcomes
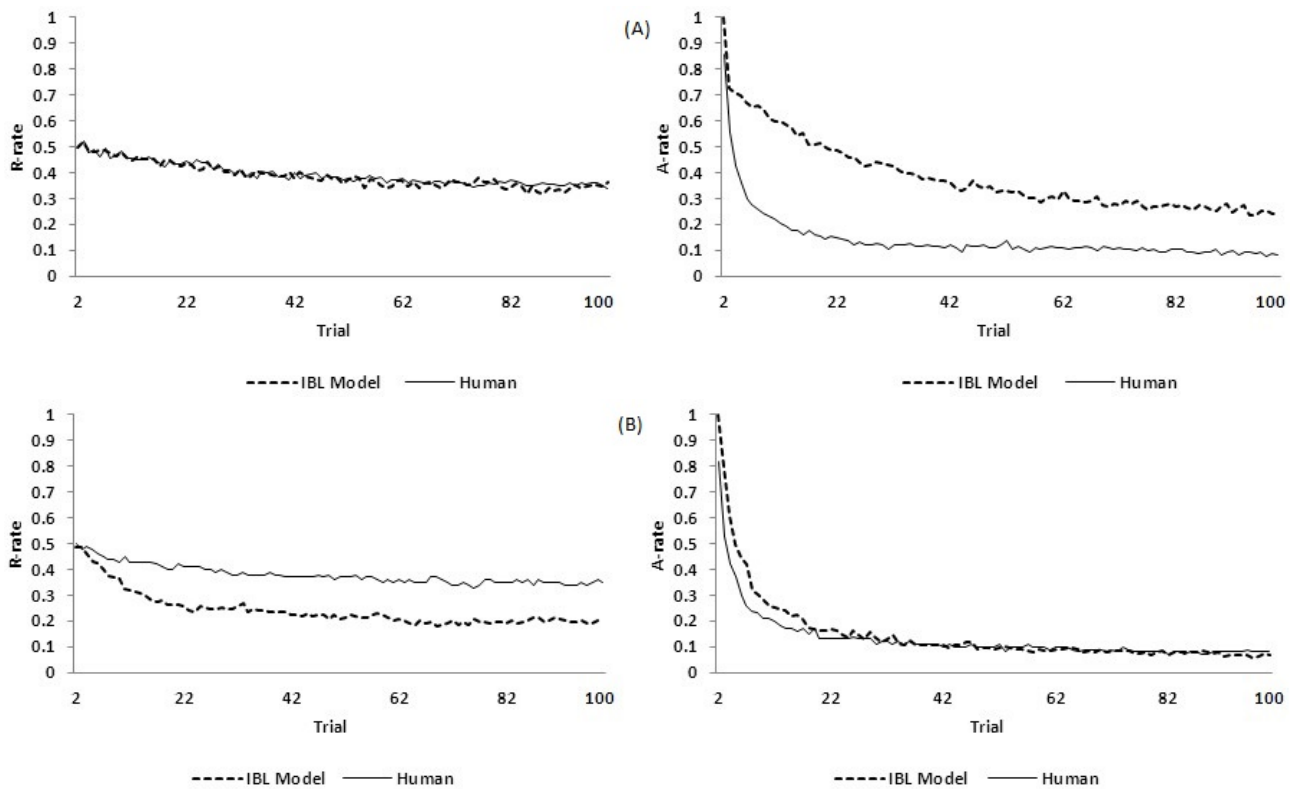
Figure 5. The generalization of the IBL model in the TPT's competition set. (A) The model's parameters were calibrated on the R-rate per problem measure in the TPT's estimation set. (B) The model's parameters were calibrated on the A-rate per problem measure in the TPT's estimation set.

is important, because doing so will improve our understanding of how people maximize long-term goals through the process of sequential choices from experience. Several recent model-comparison competitions have suggested the use of different dependent measures for calibrating models without a clear motivation for choosing one measure over the other. For example, the measure of model evaluation in the TPT was solely risk-taking, i.e., decision outcomes (Erev & Barron 2005); however, the measure of evaluation in the recently concluded market-entry competition (Erev, Ert, & Roth, 2010) was a combination of risk-taking (outcome) and alternations (process). Our analysis suggests that stronger and more robust models of learning should be able to explain both the decision outcomes and the process by which these outcomes came about. Future model comparison efforts should enforce both types of measures.

In this paper, we used one IBL model to showcase the benefits of calibrating models on a process measure compared to an outcome measure. This attempt maybe limited in its ability at present as we only used one model, IBL, on two datasets. However, this attempt does showcase the wider generalizability of the theory, IBLT, which has been used in literature to derive a number of models on a number of decision tasks (please see: Gonzalez, in press; Gonzalez, 2013 for more arguments).

As part of our future research, we would like to build on our current finding by calibrating and evaluating models on both the outcome and process measures in various tasks that differ in their outcome feedback and dynamics. Also, as part of future research, we would like to consider the mutual benefits of calibrating models to both process and decision outcomes especially when there are more than two measures. It would be interesting to observe the extent to which the benefits of calibrating models to different kinds of process measures carries over to different kinds of decision outcomes. In the case there are more than two measures, one could combine multiple process and outcome measures by doing a weighted sum of mean-squared deviations calculated on these measures. One could keep weights at values such that all combining measures are weighted equally during optimization. Furthermore, it would be interesting to observe how calibrating models to the process measures carries over to the outcome measures when the calibration is done at the individual level rather than at the aggregate level. These evaluations would help extend our existing knowledge on this topic and help us explore benefits and limitations for computational models in explaining both the decision outcomes and the process through which these outcomes are reached.

Institute of Technology for making the data from the Technion Prediction Tournament available.

References

Ahn, W. Y., Busemeyer, J. R., Wagenmakers, E. J., & Stout, J. C. (2009). Comparison of decision learning models using the generalization criterion method. *Cognitive Science, 32,* 1376-1402. doi:10.1080/03640210802352992

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

Biele, G., Erev, I., & Ert, E. (2009). Learning, risk attitude and hot stoves in restless bandit problems. *Journal of Mathematical Psychology, 53*(3), 155-167. doi:10.1016/j.jmp.2008.05.006

Busemeyer, J. R. (1985). Decision making under uncertainty: A comparison of simple scalability, fixed sample, and sequential sampling models. *Journal of Experimental Psychology, 11,* 538-564. doi:10.1037/0278-7393.11.3.538

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

Busemeyer, J. R., & Wang, Y.M. (2000). Model comparison and model selections based on generalization criterion methodology. *Journal of Mathematical Psychology, 44(1),* 171–189. doi:10.1006/jmps.1999.1282

Camerer, C., & Ho, T. H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica, 67*(4), 827-874. Retrieved from: http://www.jstor.org/stable/2999459

Chen, W., Liu, S. Y., Chen, C. H., & Lee, Y. S. (2011). Bounded memory, inertia, sampling and weighting model for market entry games. *Games, 2,* 187-199. doi:10.3390/g2010187

Dember, W. N., & Fowler, F. (1958). Spontaneous alternation behavior. *Psychological Bulletin, 55,* 412–428. doi:10.1037/h0045446

Dutt, V., Yamaguchi, M., Gonzalez, C., & Proctor, R.W. (2009). An Instance-Based Learning model of stimulus-response compatibility effects in mixed location-relevant and location-irrelevant tasks. In A. Howes, D. Peebles, R. Cooper (Eds.), *9th International Conference on Cognitive Modeling – ICCM2009.* Manchester, UK. Retrieved from: http://act-r.psy.cmu.edu/wordpress/wp-content/uploads/2012/12/863paper115.pdf

Erev, I., & Barron, G. (2005). On adaptation, maximization and reinforcement learning among cognitive strategies. *Psychological Review, 112*(4), 912-931. doi:10.1037/0033-295X.112.4.912

Erev, I., Ert, E., & Roth A. E. (2010). A choice prediction competition for market entry games: An introduction. *Games, 1*(2), 117-136. doi:10.3390/g1020117

Erev, I., Ert, E., Roth, A. E., Haruvy, E., Herzog, S. M., Hau, R., Hertwig, R., Stewart, T., West, R., & Lebiere, C. (2010). A choice prediction competition: Choices from experience and from description. *Journal of Behavioral Decision Making, 23*(1), 15-47. doi:10.1002/bdm.683

Erev, I., & Haruvy, E. (2005). Generality, Repetition, and the Role of Descriptive Learning Models. *Journal of Mathematical Psychology, 49*(5), 357-371. doi:10.1016/j.jmp.2005.06.009

Estes, W. K. (1962). Learning theory. *Annual Review of Psychology, 13,* 107-144. doi:10.1146/annurev.ps.13.020162.000543

Gonzalez, C. (in press). Decision Making: A Cognitive Science Perspective. Chapter 6 (pp. TBD). In Chipman, S. (Ed.), *The Oxford Handbook of Cognitive Science.* New York, NY: Oxford University Press.

Gonzalez, C. (2013). The boundaries of Instance-Based Learning Theory for explaining decisions from experience. Chapter 5, pp. 73-98. In Pammi and Srinivasan (Eds.), *Decision Making: Neural and Behavioural Approaches. Vol. 202, Progress in Brain Research.* New York, NY: Elsevier.

Gonzalez, C., Best, B. J., Healy, A. F., Bourne, L. E., Jr, & Kole, J. A. (2010). A cognitive modeling account of simultaneous learning and fatigue effects. *Journal of Cognitive Systems Research, 12*(1), 19-32. doi:10.1016/j.cogsys.2010.06.004

Gonzalez, C., & Dutt, V. (2011). Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological Review, 118,* 523-551. doi:10.1037/a0024558

Gonzalez, C., Dutt, V., & Lejarraga, T. (2011). A loser can be a winner: Comparison of two instance-based learning models in a market entry competition. *Games, 2*(1), 136-162. 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 real-time dynamic decision making. *Cognitive Science. 27*(4), 591-635. doi:10.1016/S0364-0213(03)00031-4

Green, L., Price, P. C., & Hamburger, M. E. (1995). Prisoner's dilemma and the pigeon: Control by immediate consequences. *Journal of Experimental Analytical Behaviour, 64,* 1–17. doi:10.1901/jeab.1995.64-1

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

Hills, T. T., & Hertwig, R. (2010). Information search in decisions from experience: Do our patterns of sampling foreshadow our decisions? *Psychological Science, 21*(12), 1787-1792. doi:10.1177/0956797610387443

Johnson, E. J., Schulte-Mecklenbeck, M., & Willemsen, M. (2008). Process Models deserve Process Data: Comment on Brand-

stätter, Gigerenzer, & Hertwig (2006). *Psychological Review, 115*(1), 263-272. doi:10.1037/0033-295X.115.1.263

Lebiere, C. (1999). *Blending: An ACT-R mechanism for aggregate retrievals.* Paper presented at the Sixth Annual ACT-R Workshop at George Mason University. Retrieved from: http://act-r.psy.cmu.edu/wordpress/wp-content/themes/ACT-R/workshops/1999/talks/blending.pdf

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). Oxford, UK: Psychology Press. Retrieved from: http://repository.cmu.edu/cgi/viewcontent.cgi?article=1083&context=sds

Lebiere, C., Gonzalez, C., & Warwick, W. (2009). A comparative approach to understanding general intelligence: Predicting cognitive performance in an open-ended dynamic task. In Goertzel, B., Hitzler, P., & Hutter, M., (Eds.), *Proceedings of the Second Conference on Artificial General Intelligence,* 103-107. Amsterdam-Paris: Atlantis Press. doi:10.2991/agi.2009.2

Lee, M. D., Zhang, S., Munro, M., & Steyvers, M. (2011). Psychological models of human and optimal performance in bandit problems. *Cognitive Systems Research, 12,* 164-174. doi:10.1016/j.cogsys.2010.07.007

Lejarraga, T., Dutt, V., & Gonzalez, C. (2012). Instance-based learning: A general model of repeated binary choice. *Journal of Behavioral Decision Making, 25*(2), 143-153. doi:10.1002/bdm.722

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). Mahwah, NJ: Erlbaum. Retrieved from: http://repository.cmu.edu/cgi/viewcontent.cgi?article=1087&context=sds

Nevo, I., & Erev, I. (2012). On surprise, change, and the effect of recent outcomes. *Frontiers in Psychology, 3,* 1-9. doi:10.3389/fpsyg.2012.00024

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108. doi:10.1037/0033-295X.85.2.59

Ratcliff, R., & Smith, P. (2004). A comparison of sequential sampling modles for two-choice reaction time. *Psychological Review, 111,* 333–367. doi:10.1037/0033-295X.111.2.333

Rapoport, A., & Budescu, D.V. (1992). Generation of random series in two-person strictly competitive games. *Journal of Experimental Psychology: General, 121,* 352–363. doi:10.1037/0096-3445.121.3.352

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*(1), 31-49. doi:10.1006/obhd.1996.2670

Scheres, A. & Sanfey, A.G. (2006). Individual differences in decision-making: drive and reward responsiveness affects strategic bargaining in economic games. *Behavioral and Brain Functions, 2,* 35. doi:10.1186/1744-9081-2-35

Suppes, P., & Atkinson, R.C. (1959). Markov Learning Models for Multiperson Situations, I. The Theory. *Technical Report Prepared under Contract Nonr 255(17)(NR 171-034), 21,* 1-78. Retrieved from: http://suppes-corpus.stanford.edu/techreports/IMSSS_21.pdf

Tolman, E.C. (1925). Purpose and cognition: The determiners of animal learning. *Psychological Review, 32,* 285–297. doi:10.1037/h0072784

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk Uncertainty, 9,* 195–230. doi:10.1007/BF00122574

Vandekerckhove, J., & Tuerlinckx, F. (2007). Fitting the Ratcliff diffusion model to experimental data. *Psychonomic Bulletin & Review, 14,* 1011-1026. doi:10.3758/PBR.15.6.1229

Appendix

Decision Rule

A choice is made in the model in trial t+1 as the selection of an alternative with the highest blended value as per Equation 1 (below).

Blending and activation mechanisms

The blended value of alternative j is defined as

$$V_j = \sum_{i=1}^{n} p_i x_i \qquad (1)$$

Where x_i is the value of the observed outcome in the outcome slot of an instance i corresponding to the alternative j, and p_i is the probability of that instance's retrieval from memory (for the case of our binary-choice task in the experience condition, the value of j in Equation 1 could be either Risky or Safe). The blended value of an alternative is the sum of all observed outcomes x_i in the outcome slot of corresponding instances, weighted by the instances' probability of retrieval.

Probability of Retrieving Instances

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 alternative, given by

$$P_{i,t} = \frac{e^{\frac{A_{i,t}}{\pi}}}{\sum_j e^{\frac{A_{i,t}}{\pi}}} \qquad (2)$$

Where π is random noise defined as $s \times \sqrt{2}$ s and is a free noise parameter. The noise parameter s captures the imprecision of retrieving instances from memory.

Activation of Instances

The activation of each instance in memory depends upon the *activation* mechanism originally proposed in ACT-R [2]. According to this mechanism, for each trial t, *activation* $A_{i,t}$ of instance i is:

$$A_{i,t} = \ln\left(\sum_{t_i \in 1,...,t-1} (t - t_i)^{-d}\right) + s \times \ln\left(\frac{1 - y_{i,t}}{y_{i,t}}\right) \quad (3)$$

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 or reliance on recency.

Noise in Activation

The $y_{i,t}$ term is a random draw from a uniform distribution $U(0, 1)$, and the $s \times \ln\left(\frac{1-y_{i,t}}{y_{i,t}}\right)$ term represents Gaussian noise important for capturing the variability of human behavior.

Pre-populated Instances in Memory

For the first trial, the IBL model does not have any instances in memory from which to calculate blended values. Therefore, the model is made to make a selection between instances that are pre-populated in memory. Lejarraga, Dutt, and Gonzalez [23] used a value of +30 in the outcome slot of the two alternatives' instances. The +30 value is arbitrary, but most importantly, it is greater than any possible outcomes in the TPT problems and will trigger an initial exploration of the two alternatives. We use these pre-populated values in the model in this paper.