Computational & Mathematical Organization Theory, 8, 31–47, 2002

Is It Better to Forget? Stimulus-Response, Prediction, and the Weight of Past Experience in a Fast-Paced Bargaining Task

FAISON P. GIBSON
University of Michigan Business School, 701 Tappan Street, Ann Arbor, MI 48109-1234, USA
email: fpgibson@umich.edu

Abstract

Decision makers in dynamic environments such as air traffic control, firefighting, and call center operations adapt in real-time using outcome feedback. Understanding this adaptation is important for influencing and improving the decisions made. Recently, stimulus-response (S-R) learning models have been proposed as explanations for decision makers' adaptation. S-R models hypothesize that decision makers choose an action option based on their anticipation of its success. Decision makers learn by accumulating evidence over action options and combining that evidence with prior expectations. This study examines a standard S-R model and a simple variation of this model, in which past experience may receive an extremely low weight, as explanations for decision makers' adaptation in an evolving Internet-based bargaining environment. In Experiment 1, decision makers are taught to predict behavior in a bargaining task that follows rules that may be the opposite of, congruent to, or unrelated to a second task in which they must choose the deal terms they will offer. Both models provide a good account of the prediction task. However, only the second model, in which decision makers heavily discount all but the most recent past experience, provides a good account of subsequent behavior in the second task. To test whether Experiment 1 artificially related choice behavior and prediction, a second experiment examines both models' predictions concerning the effects of bargaining experience on subsequent prediction. In this study, decision models where long-term experience plays a dominating role do not appear to provide adequate explanations of decision makers' adaptation to their opponent's changing response behavior.

Keywords: dynamic decision making, game theory, stimuls-response, reinforcement learning

Decision makers in dynamic environments use feedback from their previous decisions to adapt to an evolving situation under time pressure. For instance, experienced fire chiefs continuously adapt their attempts to control a fire as results from previous attempts reveal more about it (Klein et al., 1993). Air traffic controllers and police dispatchers monitor the results of previous allocation decisions to determine how they might alter their course of action (Joslyn and Hunt, 1998; Kanfer and Ackerman, 1989). During calls to delinquent debtors, telephone-based credit collectors weigh responses to their offers in adjusting their next offer up or down (Gibson and Fichman, 1998; Sutton, 1991). Understanding how decision makers adapt using feedback is important because it provides a lever for influencing and improving their decision making.

Brehmer and others have suggested that an important determinant of decision makers' ability to adapt from feedback is the internal causal model they develop of the task (Brehmer, 1990, 1992, 1995; Dienes and Fahey, 1995; Gibson et al., 1997; Gibson, 2000; Kleinmuntz, 1993; Sterman, 1994). This observation suggests that more experienced decision makers with better developed models of how causes lead to events should perform better in evolving environments, and they do (Klein et al., 1993). It also suggests that a way to improve decision makers' performance is to improve their causal model of the environment. However, rapid adaptation to feedback is not limited to experienced decision makers, and it appears not to require in-depth knowledge about the causal structure of the task. Paich and Sterman (1993) observe that decision makers in a complex market task adapt their decisions so as to pursue more successful strategies although their level of performance does not appear to be based on a sophisticated understanding of cause and effect in the task. Gibson et al. (1997) similarly observe that decision makers use feedback about previous decision outcomes to adapt their decisions to changed context in a production management task, even though the decision makers do not display a systematic understanding of causal factors within the task environment.

Recently, several authors have proposed stimulus-response (S-R) learning to provide an account of how decision makers rapidly form simple internal models that associate actions with rewards without necessarily taking into account causation (Dienes and Fahey, 1995; Erev and Roth, 1998; Fudenberg and Levine, 1998; Roth and Erev, 1995). A frequent base assumption in these models is that the decision maker is in a stable environment that does not alter over time (Fudenberg and Levine, 1998, p. 31). However, dynamic environments do not display stable response tendencies thereby leading to the question of whether these models can account for behavior in dynamic environments.[1]

This study examines base (long-memory) S-R models and a simple variation of these models (short-memory), in which past experience is rapidly forgotten, as explanations for decision makers' adaptation in an evolving Internet-based ultimatum bargaining task. The next section reviews the psychological assumptions and mechanics of S-R models as applied to dynamic tasks. In Experiment 1, both long-memory and short-memory models are used to provide accounts of decision makers' behavior when they are taught to predict their opponents' responses in a bargaining task. This task follows rules that may be the opposite of, congruent to, or unrelated to a second task in which decision makers must choose the deal terms they will offer. Both models provide a good account of the prediction task, but only the short-memory model begins to account for subsequent behavior in the bargaining task. To determine how dependent the effects observed in Experiment 1 are on the prediction training corpus used in that experiment, Experiment 2 examines the effects of bargaining experience with opponents using different response rules on decision makers' subsequent predictions of their opponents reactions. Decision models where long-term experience plays a dominating role in determining behavior do not appear to provide an adequate explanation of behavior in the very simple environment examined in this study.

S-R Learning and Dynamic Tasks

S-R learning encompasses a broad class of models that are currently being investigated in repeated decision making (Dienes and Fahey, 1995; Erev and Roth, 1998; Fudenberg and Levine, 1998; Rapoport et al., 1997; Roth and Erev, 1995), skill automatization (Logan, 1988, 1990, 1992), binary categorization (Erev, 1998), and animal learning (Barto et al., 1989). As accounts of decision making, S-R models assume that decision makers represent choices as a set of discrete options, a_i, one per possible action, for which evidence accumulates in the course of learning. For instance, the action options for a bargainer who can make high, medium, or low bids are just high, medium, and low. At any point in time when called upon to take action, the decision maker chooses an action based on the weight, w_i, of evidence that has accumulated for that action relative to other actions given by:

$$Evidence(a_i) = \frac{w_i}{\sum_{j=1}^{k} w_j} \tag{1}$$

where k is the total number of available action options, three in the case of the example.

Equation (1) might be thought of as giving the decision maker's anticipation of success using action a_i relative to other available options. For instance, if the bargainer has had success with low bids nine times and medium or high bids one time each, the weight of evidence for the low bid is nine-elevenths. The decision maker is assumed to use this expectation or judgment to make a choice by either choosing the action with highest expectation (Fudenberg and Levine, 1998, best response) or, as assumed in this study, choosing stochastically with the probability of each action given by the relative weight of evidence in its favor (Erev and Roth, 1998). In the first case, the decision maker only makes the offer with the highest expectation of success, and, in the latter, the decision maker engages in probability matching, a frequently noted pattern in choice under uncertainty (e.g., Erev et al., 1999).

An important feature of this decision rule is that it is essentially associative. Options are chosen based on their associations with past successes, not based on a theory the decision maker may possess of how his or her actions cause outcomes. There is much evidence from more complex environments compatible with the assumption that developing an explicit theory of cause and effect in the environment is not a prerequisite for performance improvement. Sterman (1989) observes that decision makers improve with versions of his stock management task but are not able to immediately transfer these performance improvements when some of the parameters of the task change. Brehmer (1995) makes similar observations for subjects playing the role of firefighters. Finally, Stanley et al. (1989) observe that only 15% of decision makers were able to state a rule that properly took into account causality as they learned to manage a simulated factory, although performance improvement across decision makers was statistically significant. More significantly, in this study, decision makers who were able to state a useful causal model did so only after showing significant performance improvement.

Evidence for each decision option accumulates based on the perceived reward, r_t, that would have been received at time t as a result of taking action a_i or of *predicting* that a_i would succeed (for an elaboration of this predictive view, see Barto et al., 1989; Sutton, 1988). The weight of evidence for each option increases according to the following equation:

$$w_{i,t+1} = \begin{cases} w_{i,t} + r_t, & \text{if selected and success} \\ w_{i,t} + r_t/(k-1), & \text{if not selected and selected option fails} \\ w_{i,t}, & \text{otherwise} \end{cases} \tag{2}$$

where $k-1$ is the total number of decision options less the selected option. Note from the equation that the decision maker increases the weight of unchosen alternatives if the chosen option fails. This assumption is quite common in modeling decision behavior (Fudenberg and Levine, 1998, pp. 120–121).[2]

There are different possible interpretations of r_t that involve assumptions about how decision makers perceive rewards. The work in this study uses two assumptions for r_t. The long-memory model assumes that decision makers' perception of r_t is equal to *a fixed fraction, α_1, of the decision maker's initial expectations as measured at the beginning of the experiment* (i.e., $r_t = \alpha_1 \sum_{j=1}^{k} w_{j,t=0}$). This same core assumption has been adopted by Erev and his collaborators as a point of departure in using S-R models to account for how decision makers learn in games (e.g., Erev et al., 1999; Erev and Roth, 1998; Roth and Erev, 1995) and also underlies a number of related learning theories (Fudenberg and Levine, 1998. p. 31).

Note that the constant learning rate in Eq. (2) assumes that decision makers already give strong weight to their initial expectations at the start of learning. As learning progresses using Eq. (2), the ability of new evidence to alter decision maker behavior decreases to the point where it becomes infinitesimally small, and decision makers cease adapting (Fudenberg and Levine, 1998). This feature of the model runs counter to evidence from dynamic tasks where decision makers are able to more rapidly adjust their behavior to changed environmental circumstances. For instance, Gibson et al. (1997) present evidence that decision makers experience an initial performance deficit when asked to manage a simulated factory to a new production goal but rapidly adapt.

The short-memory model is an alternative that allows decision makers to adapt. The short-memory model assumes that decision makers' perception of r_t is equal to *a fixed fraction, α_2, of the decision maker's prior expectations as accumulated in the task up to the point of the particular decision* (i.e., $r_t = \alpha_2 \sum_{j=1}^{k} w_{j,t}$). Thus, the key differentiating feature of this model is that decision makers continuously discount prior experience in light of new evidence, also an assumption in previous adjustments to S-R learning where a "forgetting" parameter is included to improve fit in stable environments (Erev and Roth, 1998; Fudenberg and Levine, 1998; Roth and Erev, 1995). With sufficient experience, adaptive decision makers forget the past.

As described, a clear implication of S-R learning is that teaching decision makers to predict rewards should influence their subsequent selection of action options in predictable ways. As indicated by the construction of the short-memory and long-memory models, the strength and duration of any experimentally induced bias should depend on the weight the decision maker places on past evidence. These issues are examined in detail in Experiments 1 and 2 below.

Experiment 1: Prediction and Bargaining

This experiment used a fast-paced bargaining task, the Collections task, to test both short-memory and long-memory accounts of S-R learning as explanations of decision makers' behavior when they learn to predict an opponent's responses to their offers and then must actually bargain.

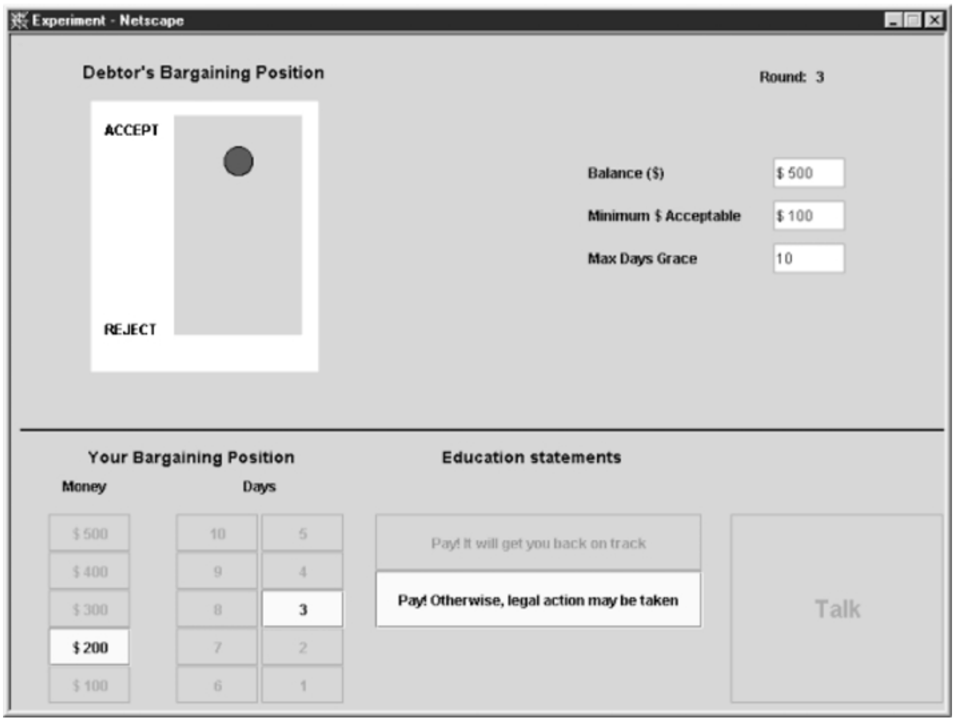

Figure 1. Bargaining screen.

Task

The Collections task, shown in figure 1, is a repeated bargaining game between debtors and collectors in which collectors make offers to debtors, and debtors respond by either accepting or rejecting. Subjects play the role of the credit collectors. As shown in figure 1, during each contact, the subject makes an offer to the debtor within a time limit of four seconds.[3] Subjects' goal is to get the debtor to agree to as high a payment in as few days as possible. To make their offers, subjects click on a dollar amount (between $100 and $500), a number of days (between one and ten) in which the dollar amount must be paid, and an education statement (positive:[4] "Pay, it will get you back on track!"; negative: "Pay, otherwise legal action may be taken!").[5] Once subjects have completed their selections, they click on the button labeled *Talk*. A male voice responds only with the statements "Accept" or "Reject".

The collections task is hosted on the Internet, and subjects are told that they are bargaining with debtors on the other side of a movable wall that divides the room where the experiment takes place. Subjects are also instructed that, if debtors do not accept their offers, they should compromise by offering to wait more days for payment, accepting smaller payments, or some combination of the two. Subjects are further instructed that they should try the different education statements and determine for themselves which are more effective.

In reality, the debtor's responses are based on whether the subject (collector) has chosen the appropriate education statement as follows:

$$Response = \begin{cases} \text{Accept with 85\% probability,} & Bargainer\_Education = Accept\_type \\ \text{Reject with 85\% probability,} & Bargainer\_Education \neq Accept\_type \end{cases} \tag{3}$$

where *Bargainer_Education* is the education statement selected by the subject (collector) and *Accept_type* is the type of education that the debtor responds to. Responses are probablistic to add uncertainty to the task with 85% chosen to provide a high enough level of validity that subjects will not have too much trouble discerning it in the time allotted for the experiments (Naylor and Clark, 1968). No other considerations, such as the days or dollars components of the offers, enter into the debtor's decision to accept. While this structure is very simple, it corresponds to decision makers' intuitive representation of targets as requiring harder and softer approaches in such functioning environments as credit collection call centers and police interrogation centers (Rafaeli and Sutton, 1991; Sutton, 1991).

Method

Forty-eight University of Michigan undergraduate business students participated in the study for $10 pay. As detailed in Table 1, subjects sequentially completed four activities in the experiment: initial expectations, prediction training, bargaining, and factual questions. The experimental manipulation occurred in the second activity, prediction training, where subjects were randomly assigned in equal groups to predict the responses of debtors using control, negative, or positive response rules.

The first activity, initial expectations, obtained baseline data on subjects' beliefs concerning the effectiveness of positive and negative education prior to the experimental manipulation. In prediction training, debtors using the positive response rule accepted only after positive education, those using the negative response rule only after negative education, and those using the control response rule accepted according to each of these rules one-half the time. In the bargaining task, subjects made their own offers consisting of dollars, days, and education. In this task, all debtors responded favorably to positive education based on Eq. (3). After bargaining, in the factual questions, subjects indicated what they considered to be the single most important element in determining whether the debtor would accept or reject their offer.

Reinforcement Learning Model Assumptions. The implementation of the short-memory and long-memory models made two important psychological assumptions about decision maker behavior beyond those already addressed. The first concerned how decision makers represented the offers as they learned in the task. The simulations in this experiment assumed that evidence was accumulated for each offer component separately and independently. The primary justification for this assumption is that it was suggested to decision makers by the instructions. Second, given that there were 100 separate combinations of offer components,[6] decision makers accumulating evidence over combinations of components would accumulate evidence for any given combination only very slowly. If decision makers

Table 1. Activities by experimental session. All activities used the collections task.

Activity	Description
Initial expectations	Subjects completed 20 prediction drills in which they had to indicate within four seconds whether the debtor would accept or reject a given offer by clicking on the graph in figure 1. Offers were presented to subjects as yellow highlighting on each of the offer's component buttons. For instance, an offer to pay \$100 in 2 days with negative education was indicated by highlighting the \$100, 2 days, and negative education buttons. Subjects then had to click on the graph to indicate a prediction of accept or reject. Subjects received no feedback regarding the correctness of their predictions. Offers were constructed as follows. There were four offers for each of the dollar amounts from \$100 to \$500. For each dollar amount, two of the offers contained positive education and two negative education. Days and dollar amounts were correlated with offers of one and two days grace coinciding with \$100, two and three days grace with \$200, etc. Education statements were orthogonal to each of these other two bid components. The order of presentation was randomized across subjects.
Prediction training	Subjects completed 50 four second prediction drills in random order receiving feedback about the correctness of their predictions depending on whether they were in the positive, negative, or control prediction conditions. The predictions were performed exactly as in the initial expectations task. The fifty offers were constructed so that each of the possible even days (2, 4, 6, 8, 10) was grouped with all of the dollar amounts and all of the education statements. Across the fifty offers, all offer components were uncorrelated. Feedback consisted of a red dot appearing in the place the subject should have clicked. In the positive condition, positive education caused a red dot to appear on accept and negative education caused a red dot to appear on reject. The negative education condition produced dots in the exact opposite pattern. The control condition responded according to the negative condition one-half the time and the positive condition one-half the time so that neither positive nor negative education was favored.
Bargaining	Subjects completed fifty randomized four-second rounds of bargaining in which they specified all components of their offers to debtors by clicking on the buttons they chose for dollars, days, and education. Debtors responded favorably to positive education and negatively to negative education using the decision rule in Eq. (3). Subjects were told to get the best deals they could because their (fictitious) boss would give them a (fictional) prize if they did the best.
Factual questions	Subjects were asked to select the offer element (dollars, days, and education) that had the most influence on debtors' decisions to accept or reject their offers.

considered components independently, then when considering what to do for education, they would only accumulate evidence over two options. Were decision makers to somehow use a combination of the two representations, the simpler representation where components were considered independently would accumulate evidence more rapidly and thereby dominate. Thus, the simulations were simplified to only considering how decision makers accumulated evidence for the education component of each offer, the part that mattered most for the offer being accepted in this task. This assumption is examined below in an analysis of subjects' responses to the factual questions.

The second assumption was how to set α_1 and α_2, the size given to evidence weight updates, r_t, in the long-memory and short-memory models respectively. In the case of α_1, a simple approach was to assume that the responses to the 20 initial expectations questions represented the subjects' initial evidence weights and that subjects updated these weights by adding one to the evidence weights of successful options as they bargain (Fudenberg and Levine, 1998, p. 31). To give initial evidence the same weight at the start of learning with short-memory models, α_2 was set so that each weight update would be one-twentieth the total accumulated evidence weight to that point in the experiment.

Results

The evolution of subjects' belief by condition that positive education would produce an accept and negative education a reject (the positive hypothesis) was of interest in the prediction task. Counts of subjects' predictions congruent with this hypothesis were summed over groups of ten contacts, producing five data points for each subject from start to finish of the task. For the long-memory and short-memory simulation models, predictions were generated by using Eq. (1) to estimate the model's average propensity to make predictions congruent with the positive hypothesis based on the evidence accumulated up to the mid-point of that group of 10. These estimates were then rounded to the nearest whole number.

Figure 2 compares the predictions of the long-memory and short-memory S-R models with subjects' performance. Both the long-memory and short-memory models predicted

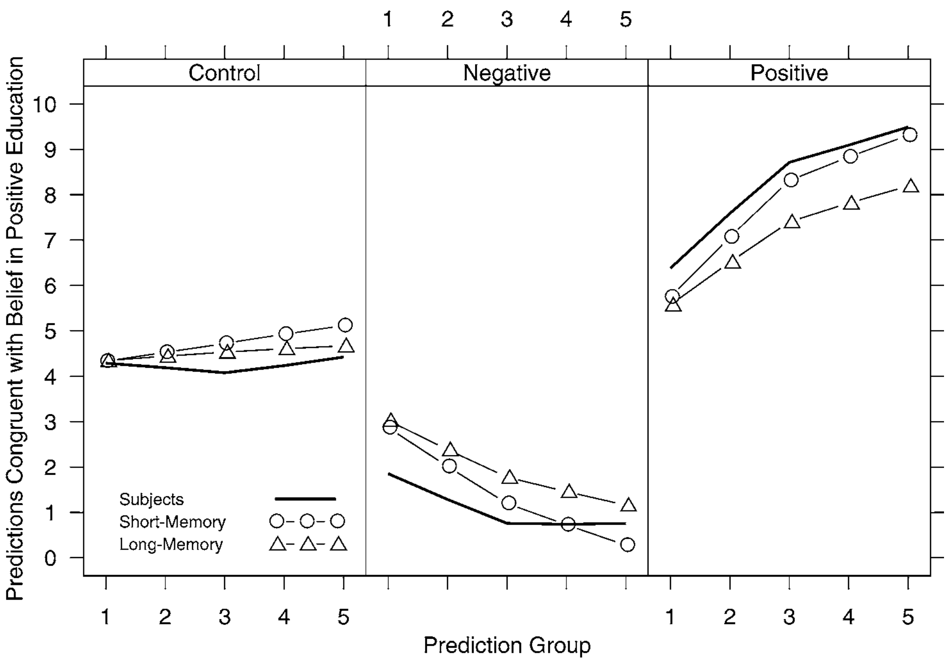

Figure 2. By prediction condition (control, negative; positive), a comparison between subjects, short-memory models and long-memory models of the number of predictions congruent with the belief that positive education leads to compliance.

Table 2. By prediction condition, difference (Δ) between number of model and subject predictions congruent with the belief that positive education is effective.

	Short-memory			Long-memory		
	Δ	t_{15}	p	Δ	t_{15}	p
Control	−2.5	−2.01	0.06	−1.5	−1.21	0.25
Positive	2	1.02	0.32	5.88	3.00	0.01
Negative	−1.75	−0.94	0.36	−4.5	−2.41	0.03

that subjects in the positive condition would increase their predictions congruent with the positive hypothesis, would decrease these predictions in the negative condition and evolve toward making predictions congruent with the positive hypothesis 50% of the time in the control condition.

Subject data confirmed both the long-memory and short-memory models' predictions concerning the direction of the effects. Across the prediction task, subjects in the positive condition made significantly more predictions congruent with the positive hypothesis than subjects in the control and negative conditions ($t_{45} = 13.48$, $p < 0.001$), and subjects in the negative condition made significantly less than those in the control condition ($t_{45} = -6.56$, $p < 0.001$). Similarly to both models, subjects in the control condition showed a modest but significant linear increase toward 50% in the number of predictions congruent with the positive hypothesis ($t_{45} = 3.23$, $p < 0.01$). Subjects in the negative condition showed a significant decrease relative to controls ($t_{45} = -2.37$, $p < 0.05$), and those in the positive condition an increase relative to controls and negative condition subjects ($t_{45} = 7.86$, $p < 0.001$).

Table 2 shows that while both long and short-memory models provided good fits to the number of subject predictions congruent with the belief that positive education was effective, the short-memory model's fits tended to be closer. The short-memory model's prediction was only marginally different from control subjects and insignficantly different for subjects in positive and negative prediction training. By counter, the long-memory model was significantly different from subjects for positive and negative prediction training.

Figure 3 compares subjects' use of positive education with that predicted by both models in the bargaining task. As in prediction training, subjects made 50 offers that were grouped by tens for analysis. The long-memory model predicted that subjects in the positive prediction condition would almost entirely use positive education, those in the negative prediction condition would initially use almost no positive education, increasing their use to a little under half the time by the end of the task, and those in the control prediction condition would use positive education about half the time at the start of the task increasing to eight out of ten by the end of the task. It also predicted that the change in performance from start to finish of the task would be highest for subjects in the negative prediction condition, next highest for controls, and lowest for subjects in the positive prediction condition. As further apparent in figure 3, the short-memory models agreed with both the initial deficits predicted by the long-memory model and the ordering in subjects improvement. However, it predicted much higher final performance for negative and control conditions.

As apparent in figure 3, human subjects results agreed with the initial performance difference and the order of performance improvement that both models predicted between

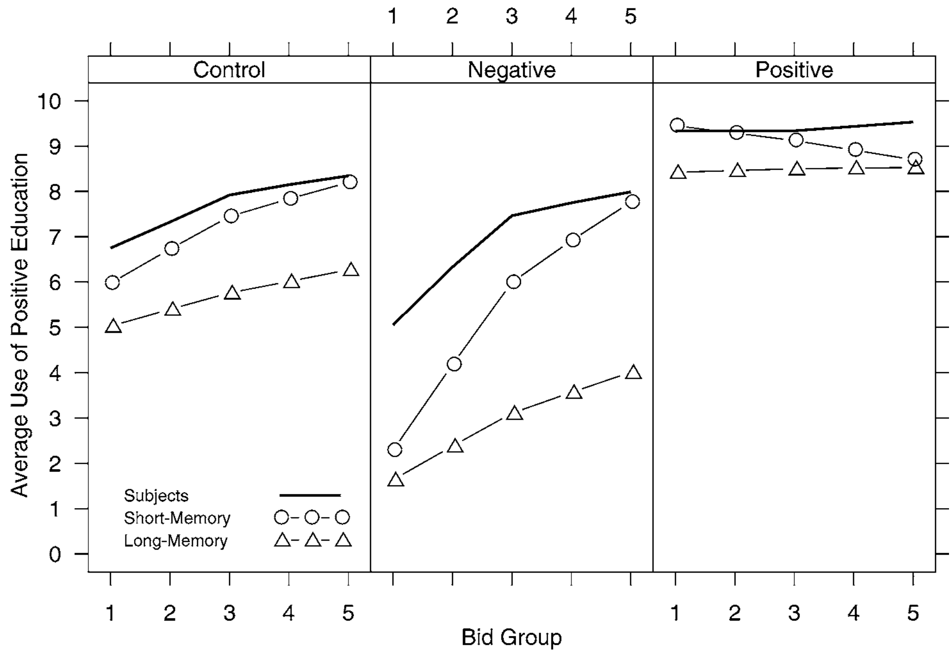

Figure 3. By prediction condition (control, negative; positive), a comparison between subjects, short-memory models and long-memory models of the number of the number of positive education statements used.

conditions. At the start of the bargaining task, subjects who had been in the negative prediction condition used significantly less positive education than control subjects ($t_{45} = -2.79$, $p < 0.01$), and positive condition subjects used significantly more positive education than both control and negative condition subjects ($t_{45} = 5.95$, $p < 0.001$). Negative condition subjects showed greater performance improvement during the task than control condition subjects, although the difference did not attain significance in a two-tailed test ($t_{45} = 1.51$, $p < 0.13$). Both control and negative condition subjects showed greater improvement during the task than positive prediction condition subjects ($t_{45} = -2.59$, $p < 0.05$).

Table 3 shows the short and long-memory models' fits to subject data. The short-memory model provides substantially better fits than the long-memory model. Overall,

Table 3. By prediction condition, difference (Δ) between number of model and subject offers using positive education.

	Short-memory			Long-memory		
	Δ	t_{15}	p	Δ	t_{15}	p
Control	2.31	0.84	0.42	10.125	3.66	0.01
Positive	1.44	1.01	0.33	4.5	3.16	0.01
Negative	7.44	2.49	0.025	20.13	6.73	0.001

its performance was not significantly different from that of human subjects in the control and positive conditions while the long-memory model's performance was always signficantly different from that of subjects. Although, the short-memory model's performance was significantly different from that of subjects overall in the negative condition, by the end of bargaining, it converged to subject performance. The long-memory model did not converge to subject performance.

A chief assumption in the simulation models was that subjects were aware of the separate component actions and could accumulate evidence for them separately. Subjects' answers to the question of which was the most influential element in getting debtors to pay provided indirect confirmation of this assumption. In the positive and negative conditions, a larger proportion of subjects than chance indicating that education was the most influential element in convincing debtors would show that these subjects understood the component role education played in the prediction and bargaining tasks. Seven-eighths of the subjects in both these conditions indicated education as the most influential element vs. seven-sixteenths in the control condition, a highly significant difference ($\chi^2_2 = 10.34$, $p < 0.01$). Thus, at least for subjects in the positive and negative prediction conditions, there is strong evidence that decision makers were aware of the effects of the separate component actions in the task.

Discussion

This experiment tested the S-R model based hypothesis that how subjects learned to predict a debtor's responses would influence how they behaved in subsequent bargaining. It also compared the predictions of two S-R formulations, long-memory and short-memory. Both formulations successfully predicted the direction in which subjects' performance would evolve during prediction training. However, the short-memory model provided substantially better fits, never significantly differing from subject performance, while the long-memory model was significantly different in the positive and negative training conditions.

In bargaining, subjects dealt with bargainers who conformed with their previous prediction training (positive) or who contradicted that training to various degrees (control and negative). Again, both models predicted the direction of effects. Subjects in the negative prediction condition performed worse in the bargaining task (used less positive education) than subjects in the control condition, and subjects in the positive prediction condition outperformed both.

Only the short-memory model predicted the degree of subjects' adaptation in the control and negative training conditions. Long-memory model performance was always significantly different from that of subjects. The short-memory model was at least more than two times closer to subject performance in all conditions and only differed signficantly from subjects in the negative condition. Even in the negative prediction condition, short-memory model performance converged to that of subjects.

An alternative model that relied on subjects ignoring their experience from the prediction task would not have predicted the differences at the start of bargaining that did exist for subjects. A component of performance in this type of dynamic task is that decision makers are able to quickly learn to ignore irrelevant past experience based on changed feedback to their actions alone. Subjects are nonetheless affected by this experience, at least in the early stages of learning. Thus, the design of this experiment, where decision makers had to quickly adapt to changing response tendencies, provided a good basis for distinguishing between multiple model candidates.

An important assumption underlying both short-memory and long-memory models was that subjects were aware of the separate effects of the component parts of the offer. Subjects in the positive and negative prediction conditions showed evidence that they were aware of the education component's separate effect because they were able to recognize it as the most influential element in getting their offers accepted.

Experiment 2: Decision Experience's Influence on Prediction

In Experiment 1, the prediction training was structured so that subjects saw each of twenty-five offers with both positive and negative education. Thus, prediction training provided both feedback *and* systematic exploration of the task environment. This exploration might have given subjects a hint about how to perform when they came to the bargaining task, over and above any prediction skills they learned.

A chief feature of S-R learning is that both prediction and choice depend on decision makers' anticipation of the success of any action given Eq. (1). Therefore, Experiment 2 tested whether bargaining experience influenced subsequent prediction, providing evidence of this link.

Task and Method

Experiment 2 used the same bargaining task as Experiment 1. Subjects started with the initial conditions task and then moved directly to bargaining. They then repeated the initial conditions task (i.e., performed a post-learning prediction test) and answered the factual questions. Subjects were in one of two conditions. In positive training, subjects bargained in a task where positive education was effective per Eq. (3) and those in negative training bargained in a task where negative education was effective. Thus, subjects received feedback in the task without being pushed to the same systematic exploration of task structure as in Experiment 1. Forty-three paid University of Michigan Business School undergraduates participated in this study.

Results and Discussion

Figure 4 shows subjects performance in the negative and positive training conditions. As is apparent in the figure, subjects in positive training used significantly more positive education than subjects in negative training ($t_{41} = 7.32$, $p < 0.001$). Further, the subjects in each condition showed opposite trends in their use of education ($t_{41} = 4.6$, $p < 0.001$).

The bargaining conditions also affected subjects' prediction in the post-learning test. Subjects in the positive bargaining condition showed a significantly higher tendency to predict that positive education would produce an accept and negative education a reject ($t_{41} = 8.42$, $p < 0.001$).

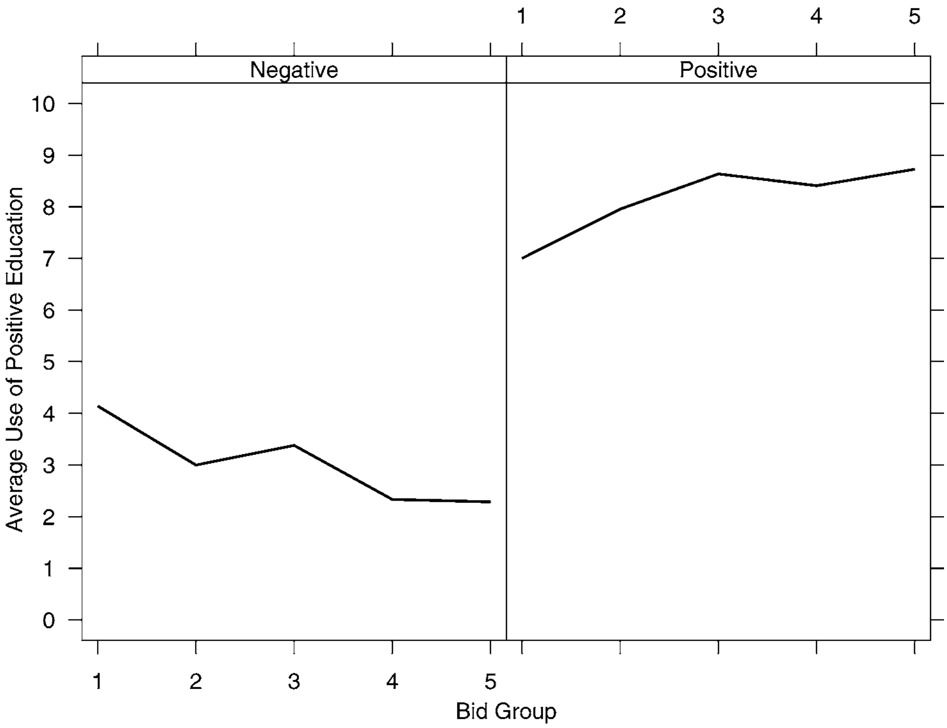

Figure 4. By bargaining condition (negative or positive), the number of positive education statements subjects used.

Finally, the positive bargaining condition allowed an additional, indirect assessment of whether subjects' participation in the control prediction training in Experiment 1 affected their subsequent bargaining performance. One possibility is that these Experiment 1 subject would significantly differ from positive bargaining subjects in Experiment 2 because they were pushed by the prediction corpus to more systematically explore the debtor's response tendencies. Comparing figures 3 and 4, subjects' performance in the positive training condition closely matched that of subjects from the control condition in Experiment 1 and was insignificantly different from it. Thus, the totality of results from this experiment suggest that the only effect prediction training had on subjects' performance in Experiment 1 was to reinforce the relevant response tendencies.

General Discussion

This study tested the predictions made by two S-R learning models, long-memory and short-memory, for how decision makers would adapt to response shifts in a very simple repeated decision environment. Experiment 1 confirmed that both models provided an adequate account of how decision makers learned to predict the debtors' responses in a preliminary prediction task, thereby confirming that S-R models may account for how decision makers anticipate effects. Experiment 1 also confirmed both models' predictions that decision makers in different prediction conditions would show different initial offer tendencies in a subsequent bargaining task as well as different rates of adaptation. However, only the short-memory model came near predicting the final level of adaptation decision makers showed when switching from predicting debtors' behavior under one response rule to bargaining with debtors who responded based on an opposite response rule.

An alternative explanation is that the structure of the prediction task gave subjects a strong hint about task structure and that subsequent behavior was not strongly tied to ability to predict. Experiment 2 demonstrated that unstructured bargaining experience influenced subjects' subsequent prediction performance, thereby more firmly establishing the link embedded in the structure of S-R models.

Implications

The models examined here have been proposed principally as explanations for behavior in stable environments that do not alter their response tendencies over time (e.g., Erev and Roth, 1998; Fudenberg and Levine, 1998; Roth and Erev, 1995). By examining the performance of these models and human decision makers in a very simple but unstable environment, this study asked to what degree decision makers conform to the type of belief updating usually assumed as the core mechanism in S-R models. The superiority of the short-memory S-R model in predicting the adaptation of decision makers who experienced opposite response functions in the prediction and bargaining tasks amplifies previous results where an additional parameter for forgetting provided marginal improvement in the models' fits to decision maker performance (e.g., Roth and Erev, 1995).

As is readily apparent from an examination of Eq. (2), the standard assumptions of S-R models predict that decision makers will be slow to react when the response tendencies of an environment are subject to change. The modeling and human subject results reported for Experiment 1 provided a case in point that this assumption may not apply to a broad class of decision environments where decision makers must adapt on the fly, even if the only basis for adaptation is outcome feedback from their decisions. A model identical in form, but where the influence of past experience quickly fades, provides a better account.

Limitations and Qualifications

The work reported here has two important limitations. First, the task is very simple. Decision makers in functioning environments typically face tasks in which they take weeks or months to gain competency (Gibson and Fichman, 1998; Joslyn and Hunt, 1998; Kanfer and Ackerman, 1989; Klein et al., 1993). However, in spite of its simplicity, the task captures the important role of education in the functioning environment, leading to the possibility that its results may generalize back to that environment. Furthermore, the task ties to a broader class of bargaining task that has been well-studied (Roth and Erev, 1995; Fudenberg and Levine, 1998; Raiffa, 1982), potentially making the results more generalizable.

Second, the modeling approach is very frugal, focusing only on one discrete memory representation for possible actions and one evidence accumulation mechanism. There is general agreement that the brain contains more than one memory and evidence accumulation mechanisms (e.g., Anderson, 1993; LeDoux, 1996; McClelland et al., 1995). However, S-R models share their assumption of discrete memory representations with theories of cognitive learning in economic games (e.g., Erev and Roth, 1998; Fudenberg and Levine, 1998), animal learning (e.g., Barto et al., 1989), specific event learning in the hippocampus (e.g., McClelland et al., 1995), and judgments of probability in MINERVA-DM (Dougherty et al., 1999). In all of these models, it is the weight of evidence provided by specific memories that drives behavior. Therefore, the reductionist approach used here speaks to a feature that is common to many models and examines its ability to provide an account of behavior in dynamic tasks.

Conclusion

This paper has presented a model which predicts that decision makers who forget will outperform those who do not in unstable environments. Empirical results with human subjects confirmed the prediction in one simple, unstable decision environment. The degree to which these results and models can be extended to other, more complex and naturalistic environments is the topic of future research.

Acknowledgment

The author would like to thank Nick Stanley, Dan Liu, and Leon Gmeindl for invaluable research assistance in performing this study. Members of the University of Michigan Decision Behavior Consortium provided invaluable assistance in critiquing the ideas in this paper.

Notes

1. Many game theorists examining learning assume stable environments, particularly when they want to understand when learning leads to long-run normative behavior (Ido Erev and Colin Camerer, personal communications). Recently, theorists have begun to formally explore organization-level adaptation in dynamic environments (Carley and Lee, 1998; Levinthal, 1997). The important question of individual adaptation to change after a period of learning, as explored in this paper, remains open.
2. This assumption is usually implemented as a decrease to the weight of the chosen option if it fails. The implementation here avoids the inconvenience of possibly negative weights.
3. This figure is derived from observations within an actual call center where collectors contacted debtors by telephone (Gibson et al., 1996).
4. In the instructions, the two education statements are indicated as meant to convey either positive or negative education.
5. In experiments, the orders of all choice options are counterbalanced.
6. Calculated as 10 possible days by 5 possible dollar amounts by two types of education.

References

Anderson, J.R. (1993), *Rules of the Mind*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Barto, A.G., R.S. Sutton and C.J.C.H. Watkins (1989), "*Learning and Sequential Decision Making*," Tech. Rep. No. COINS 89-95, Amherst: University of Massachusetts.

Brehmer, B. (1990), “Strategies in Real-Time, Dynamic Decision Making,” in R. Hogarth (Ed.) *Insights from Decision Making*. University of Chicago Press, Chicago.

Brehmer, B. (1992), “Dynamic Decision Making: Human Control of Complex Systems,” *Acta Psychologica*, 81, 211–241.

Brehmer, B. (1995), “Feedback Delays in Complex Dynamic Decision Tasks,” in P. Frensch and J. Funke (Eds.) *Complex Problem Solving: The European Perspective*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Carley, K.M. and J.-S. Lee (1998), “Dynamic Organizations: Organizational Adaptation in a Changing Environment,” in J. Baum (Ed.) *Advances in Strategic Management: Disciplinary Roots of Strategic Management Research*. JAI Press, London, Vol. 15, pp. 269–297.

Dienes, Z. and R. Fahey (1995), “Role of Specific Instances in Controlling a Dynamic System,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 1–15.

Dougherty, M.R.P., C.F. Gettys and E.E. Ogden (1999), “MINERVA-DM: A Memory Process Model for Judgments of Likelihood,” *PsyRev*, 106(1), 180–209.

Erev, I. (1998), “Signal Detection by Human Observers: A Cutoff Reinforcement Learning Model of Categorization Decisions under Uncertainty,” *Psychological Review*, 105(2), 280–298.

Erev, I., Y. Bereby-Meyer and A.E. Roth (1999), “The Effect of Adding a Constant to All Payoffs: Experimental Investigation, and Implications for Reinforcement Learning Models,” *Journal of Economic Behavior and Organization*, 39, 111–128.

Erev, I. and A.E. Roth (1998), “Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria,” *American Economic Review*, 88(4), 848–881.

Fudenberg, D. and D.K. Levine (1998), *The Theory of Learning in Games*. The MIT Press, Cambridge, MA.

Gibson, F.P. (2000), “Feedback Delays: How can Decision Makers Learn Not to Buy a New Car Every Time the Garage is Empty?” *Organizational Behavior and Human Decision Processes*, 83(1), 141–166.

Gibson, F.P. and M. Fichman (1998), “Fear Appeals: What Happens When the Credit Collector Calls?” *Unpublished Manuscript*, in revision.

Gibson, F.P., M. Fichman and D.C. Plaut (1996), “Learning in Dynamic Decision Tasks—Credit Collections at Citicorp,” Technical Report, Citicorp Behavioral Sciences Research Council.

Gibson, F.P., M. Fichman and D.C. Plaut (1997), “Learning in Dynamic Decision Tasks: Computational Model and Empirical Evidence,” *Organizational Behavior and Human Decision Processes*, 71(1), 1–35.

Joslyn, S. and E. Hunt (1998), “Evaluating Individual Differences in Response to Time-Pressure Situations,” *Journal of Experimental Psychology: Applied*, 4(1), 16–43.

Kanfer, R. and P. Ackerman (1989), “Motivation and Cognitive Abilities: An Integrative/Aptitude-Treatment Interaction Approach to Skill Acquisition,” *Journal of Applied Psychology*, 74(4), 657–690.

Klein, G.A., J. Orasanu, R. Calderwood and C.E. Zsambok (Eds.) (1993), *Decision Making in Action: Models and Methods.* Ablex Publishing Corporation, Norwood, NJ.

Kleinmuntz, D.N. (1993), “Information Processing and Misperceptions of the Implications of Feedback in Dynamics Decision Making,” *System Dynamic Reviews*, 9(3), 223–237.

LeDoux, J. (1996), *The Emotional Brain.* Simon & Schuster, New York.

Levinthal, D. (1997), “Adaptation on Rugged Landscapes,” *Management Science*, 43, 934–950.

Logan, G.D. (1988), “Toward an Instance Theory of Automatization,” *Psychological Review*, 95(4), 492–527.

Logan, G.D. (1990), “Repetition Priming and Automaticity: Common Underlying Mechanisms?” *Cognitive Psychology*, 22, 1–35.

Logan, G.D. (1992), “Shapes of Reaction-Time Distributions and Shapes of Learning Curves: A Test of the Instance Theory of Automaticity,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(5), 883–914.

McClelland, J.L., B.L. McNaughton and R.C. O’Reilly (1995), “Why There are Complementary Learning Systems in the Hippocampus and Neocortex: Insights from the Successes and Failures of Connectionist Models of Learning and Memory,” *Psychological Review*.

Naylor, J.C. and R.D. Clark (1968), “Intuitive Inference Strategies in Interval Learning Tasks as a Function of Validity Magnitude and Sign,” *Organizational Behavior and Human Performance*, 3, 378–399.

Paich, M. and J.D. Sterman (1993), “Boom, Bust, and Failures to Learn in Experimental Markets,” *Management Science*, 39(12), 1439–1458.

Rafaeli, A. and R.I. Sutton (1991), "Emotional Contrast Strategies as Means of Social Influence: Lessons from Criminal Interrogators and Bill Collectors," *Academy of Management Journal*, 34(4), 749–775.

Raiffa, H. (1982), *The Art and Science of Negotiation*. Harvard University Press, Cambridge, MA.

Rapoport, A., I. Erev, E.V. Abraham and D.E. Olson (1997), "Randomization and Adaptive Learning in a Simplified Poker Game," *Organizational Behavior and Human Decision Processes*, 69(1), 31–49.

Roth, A.E. and I. Erev (1995), "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 8, 164–212.

Stanley, W.B., R.C. Mathews, R.R. Buss and S. Kotler-Cope (1989), "Insight without Awareness: On the Interaction of Verbalization, Instruction, and Practice in a Simulated Process Control Task," *Quarterly Journal of Experimental Psychology*, 41A(3), 553–577.

Sterman, J.D. (1989), "Misperceptions of Feedback in Dynamic Decision Making," *Organizational Behavior and Human Decision Processes*, 43, 301–335.

Sterman, J.D. (1994), "Learning in and about Complex Systems," *System Dynamics Review*, 10(2/3), 291–330.

Sutton, R.I. (1991), "Maintaining Norms about Expressed Emotions: The Case of Bill Collectors," *Administrative Science Quarterly*, 36, 245–268.

Sutton, R.S. (1988), "Learning to Predict by the Method of Temporal Differences," *Machine Learning*, 3, 9–44.

Faison P. Gibson is an assistant professor at the University of Michigan Business School. He holds a Ph.D. from Carnegie-Mellon University.