Faculty Research

University
of Michigan
Business
School

Learning and Transfer in Dynamic
Decision Environments

Faison P. Gibson
University of Michigan Business School

Learning and transfer in dynamic decision environments

Faison P. Gibson

University of Michigan Business School

701 Tappan Street

Ann Arbor, MI 48109-1234

`fpgibson@umich.edu`

May 24, 2002

Abstract

An important aspect of learning is the ability to transfer knowledge to new contexts. However, in dynamic decision tasks, such as bargaining, firefighting, and process control, where decision makers must make repeated decisions under time pressure and outcome feedback may relate to any of a number of decisions, such transfer has proven elusive. This paper proposes a two-stage connectionist model which hypothesizes that decision makers learn to identify categories of evidence requiring similar decisions as they perform in dynamic environments. The model suggests conditions under which decision makers will be able to use this ability to help them in novel situations. These predictions are compared against those of a one-stage decision model that does not learn evidence categories, as is common in many current theories of repeated decision making. Both models' predictions are then tested against the performance of decision makers in an Internet bargaining task. Both models correctly predict aspects of decision makers' learning under different interventions. The two-stage model provides closer fits to decision maker performance in a new, related bargaining task and accounts for important features of higher-performing decision makers' learning. Although frequently omitted in recent accounts of repeated decision making, the processes of evidence category formation described by the two-stage model appear critical in understanding the extent to which decision makers learn from feedback in dynamic tasks.

1 Introduction

Dynamic decision tasks such as firefighting, process control, and bargaining, require repeated decisions under time pressure. A common difficulty for learning in these tasks is that outcomes are frequently the result of sequences of decisions, not just one decision in isolation. For instance, a bargainer who starts with a very high asking price that is rejected may then be able to get a lower price accepted. If the bargainer is skilled, the lower price will still be higher than what would have been accepted as an initial offer. The price improvement is not due to either one of the asking prices in isolation but to their sequence (Cialdini, 1984). Experimental investigations of decision makers in tasks with similar and more complex sequential dependencies indicate that decision makers show performance improvement (e.g., Diehl & Sterman, 1995; Gibson, 2000), but in the time allotted, they are less able to develop knowledge about the task that they can apply to situations they have not yet seen. With practice, our example bargainer might get better at manipulating sequences of prices on a specific item within a narrow range, but she will have difficulty applying this skill to other items or price ranges.

Can theory provide guidance in designing a decision environment to help this bargainer? Poor ability in transferring knowledge between different task contexts is so pervasive in repeated decision making that many theories suffice with the assumption that decision makers learn based on success and failure given the specific evidence at the time of the decision (e.g., Dienes & Fahey, 1995; Fudenberg & Levine, 1998; Roth & Erev, 1995), effectively ruling out the possibility of transfer (Logan, 1988). However, performance in functioning dynamic environments suggests that expert decision makers routinely and successfully apply their knowledge to novel situations (Kanfer & Ackerman, 1989; Klein, Orasanu, Calderwood, & Zsambok, 1993; Joslyn & Hunt, 1998). Experienced air traffic controllers, firefighters, and police dispatchers are all able to perform more effectively in novel situations than less experienced decision makers.

This contrast in theory and results suggests that the initial question can be further refined:

(1) Under what conditions does theory predict that experienced decision makers are able to transfer their knowledge to novel situations? (2) What different types of information does theory predict will influence novice decision makers? To help address these questions, this paper develops a two-stage connectionist model of learning in dynamic tasks. In stage one, decision makers form internal categorizations of evidence available in the task environment. Then, in stage two, they use these internal categorizations to cue a decision. As decision makers better develop the ability to categorize evidence, the two-stage model predicts that decision makers will become better at transfer to outwardly novel situations where the same categories apply.

The next section reviews prior work in dynamic decision making to help constrain construction of the model. After that, the model is elaborated with comparison to a one-stage model that learns based on success or failure in light of specific evidence, with no internal categorization, as is frequently assumed in studies of repeated decision making. Then both models are instantiated in a simulation experiment to make predictions for human learning and transfer in an Internet-based bargaining task. Simulation modeling is well-suited to comparing theories of performance in dynamic environments because it leads to quantitative predictions that facilitate the identification of aspects of theory that may and may not conform to actual behavior (Gobet & Simon, 2000). Both models correctly predict the effectiveness of knowledge supplied to affect naïve decision makers. The two-stage model provides closer fits and provides a more accurate account of decision behavior in a new task where identical categories apply to evidence presented in a novel sequence.

2   Prior Results on Learning and Transfer

With experience, decision makers may develop an ability to transfer that is limited by: (1) how closely new contexts resemble those already encountered; and (2) the length of sequential dependency between decisions in the task. In a process control task, Gibson, Fichman, and Plaut's (1997) subjects were able to show performance approximating their training performance

when asked to achieve goals that differed by no more than one production level from what they had previously experienced. Beyond that range, initial performance for the transfer goal degraded, although subjects adapted more quickly to the new goals than they had when first learning the task. In the same task with longer sequential dependencies between decisions, Gibson's (2000) subjects also showed better performance near goals they had previously encountered. However, their overall performance on near and far goals combined was no better than what they had achieved during initial training.

An apparent explanation for these results is provided by Sterman and his collaborators (Diehl & Sterman, 1995; Paich & Sterman, 1993; Sterman, 1989a, 1989b). In particular, Diehl and Sterman (1995) have demonstrated that although decision makers show performance improvement with practice, they fail to develop heuristics that account for the sequential dependencies in the tasks that can help them transfer to new contexts. This deficit becomes only more apparent as the length of the sequential dependencies increases.

However, Stanley, Mathews, Buss, and Kotler-Cope (1989, Experiment 4) show that, after extensive practice over multiple experimental sessions, high performing decision makers in a task with short sequential dependencies were able to produce rules that, when provided to novices, improved their performance. While these results do not report the extent to which the rules captured sequential dependencies, they do indicate: (1) that decision makers develop knowledge that captures general features of the task; and that (2) this general knowledge can be communicated to novices to improve their performance.

Given the difficulty of describing many dynamic environments with simple rules (e.g., Sterman, 1989b), it would be beneficial if simply displaying diagnostic information, without the structure of a rule, could improve performance. Although one might argue that increasing the information to process would increase decision makers' workload, thereby reducing their performance (e.g., Jarvenpaa, 1989), the diagnostic benefit of the information might outweigh any such cost. In this regard, Sengupta and Abdel-Hamid (1993) show that graphically displaying the

values of variables that have a significant effect on outcomes improves decision makers'

performance, even in the presence of long sequential dependencies. Further, Gibson (2000)

demonstrates that displaying past decisions and outcomes in tasks with sequential dependencies

significantly improves performance. Thus, even without the structure inherent in a stated rule,

diagnostic information about key variables improves decision makers performance <u>across a broad</u>

<u>range of subjects</u>. The one and two-stage learning models presented next both provide an account

of how such diagnostic information can aid decision makers in dynamic environments.

<div align="center">Learning to Make Decisions Based on Evidence and Categories</div>

Figure 1 displays the one and two-stage learning models. Both models are based on a simple

analogy with the brain in which all processing is conceived as patterns of activation across layers

of neurons (McClelland, 2001).[1] As is common in studies of repeated decision making, the

models further assume that decision making in dynamic tasks consists of choosing among

competing options at each decision point based on past success and failure (e.g., Dienes & Fahey,

1995; Erev & Roth, 1998; Logan, 1988; Roth & Erev, 1995). However, the models differ on

whether specific evidence available at each decision point is used to directly cue this choice

(one-stage model) or is first classified into categories that then cue the choice (two-stage model).

We start with the one-stage model. More formally in this model, the likelihood

($\mathrm{act}(option_{i'})$) that $option_{i'}$ of m possible options will be chosen is based on the decision maker's

assessment of its likelihood of success given the cumulative support (cum_ev) it is receiving

relative to other options:

$$\mathrm{act}(option_{i'}) = \frac{e^{cum\_ev_{i'}}}{\sum_{i=1}^{m} e^{cum\_ev_i}} \tag{1}$$

After this assessment occurs, choice is a competition among alternatives using each alternative's

perceived likelihood of success as its likelihood of winning (Erev & Roth, 1998; Fudenberg &

Levine, 1998).

---

[1] Decision makers are not assumed to be explicitly aware of the calculations used to simulate these patterns of activation, a common assumption in research on cognitive performance (e.g., Simon & Gobet, 2000, note 7).

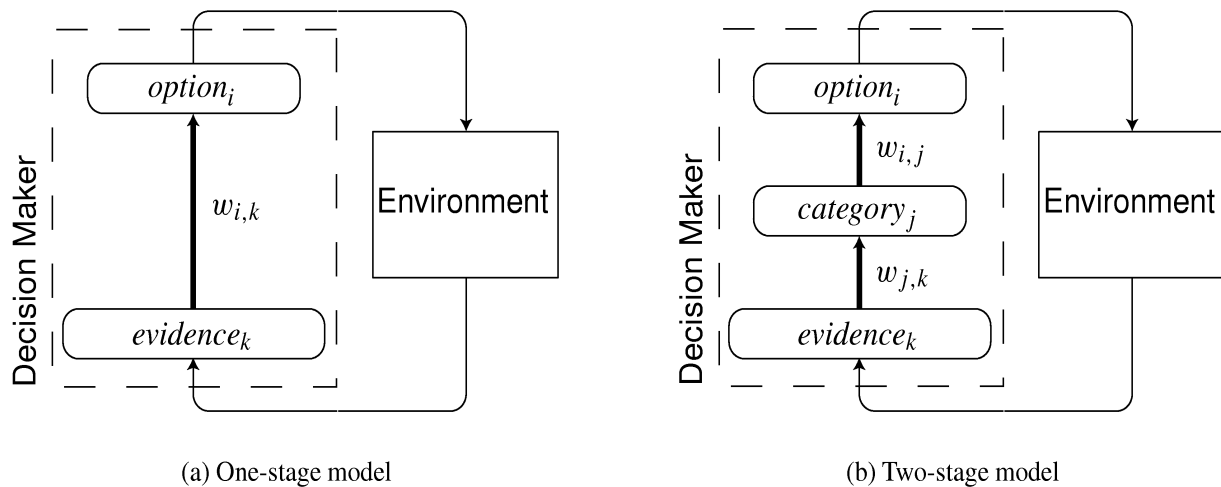(a) One-stage model   (b) Two-stage model

Figure 1: One-stage and two-stage learning models. See text for details.

The cumulative support for each decision option is calculated by summing over the absence or presence (0/1) of all n possible pieces of specific evidence ($evidence_k$) weighted ($w_{i',k}$) by their past correlation with the success of the particular decision option (Stone, 1986):

$$cum\_ev_{i'} = \sum_{k=1}^{n} evidence_k \, w_{i',k} \qquad (2)$$

Thus, the one-stage model accounts for the effect of diagnostic inputs in the environment based on the strength of their correlation with past decision outcomes as suggested by Sengupta and Abdel-Hamid (1993). Additionally, when information from past decision contexts is still relevant for the current decision, the model only accounts for this information if it is provided as part of its inputs, thereby predicting the success of Gibson's (2000) display of past information to help decision makers in a task with sequential dependencies. Further in this regard, the model accounts for difficulty in learning long sequential dependencies by the fact that observations of the full sequence will be less frequently available on its inputs than observations of sub-sequences.

The one-stage model assumes that decision makers learn by adjusting their evaluations of decision options based on success or failure. Since the evaluation of decision options is fully determined by the weights placed on the evidence, learning is accomplished by modifying these

weights. After each decision the following <u>delta</u> rule, used in a number of studies to model human learning (Rumelhart, Durbin, Golden, & Chauvin, 1995), is applied to change the weights based on decision outcome:

$$w_{i',k} = \eta \, \delta_{i'} \, evidence_k + w_{i',k} \tag{3}$$

where η is a learning parameter that can be set to determine the size of the weight change. In this rule, $\delta_{i'}$ is the difference between the imputed outcome for a decision option, $t_{i'}$, and the perceived probability that that decision option would succeed:

$$\delta_{i'} = t_{i'} - \mathrm{act}(option_{i'}) \tag{4}$$

where the imputed outcome, $t_{i'}$, is itself generated as follows:

$$t_{i'} = \begin{cases} 1 & , \quad \text{if } option_{i'} \text{ chosen and success} \\ 0 & , \quad \text{if } option_{i'} \text{ chosen and no success} \\ 0 & , \quad \text{if } option_{i'' \neq i'} \text{ chosen and success} \\ \dfrac{e^{cum\_ev_{i'}}}{\sum_{i \neq i''} e^{cum\_ev_i}} & , \quad \text{if } option_{i'' \neq i'} \text{ chosen and no success} \end{cases} \tag{5}$$

A key assumption in this estimation of $t_{i'}$ is that the decision maker believes only one decision option can succeed at each decision point (Bishop, 1996).

As just elaborated, the one-stage model's approach of learning direct relationships between environmental stimuli and the likelihood of a decision option's success works well as an approximation of learning without transfer in simple environments (e.g., Erev & Roth, 1998; Fudenberg & Levine, 1998). However, even limited decision maker transfer performance suggests the ability to recognize common features across decision contexts that the one-stage model does not possess.

To allow for this possibility, the two-stage model (Figure 1(b)) extends the one-stage model

by assuming an intermediate categorization stage. Decision makers process the evidence they observe in the environment into non-mutually exclusive categories ($category_j$) that then provide support for the different decision options. Decision makers learn to favor different decision options based on success or failure given the evidence category. Decision makers learn to categorize evidence by accumulating weight for and against each category conditioned on the category's performance in predicting the success and failure of the different decision options in light of the evidence.

More formally, the process for choosing a decision option as specified by Equation (1) in the one-stage model remains unchanged, but cum_ev_i is now calculated based on the graded applicability (act($category_j$)) of all of the p possible categories that may be used to classify the available evidence (i.e., $cum\_ev_i = \sum_{j=1}^{p} \text{act}category_j \, w_{i,j}$). In turn, the graded applicability of the category, is represented as a simple binomial function of the available evidence, ranging from 0 (inapplicable) to 1 (highly applicable; Rumelhart et al., 1995):

$$\text{act}(category_j) = \frac{1}{1 + e^{-cum\_ev_j}} \tag{6}$$

with $cum\_ev_j = \sum_{k=1}^{n} evidence_k \, w_{j,k}$.

As in the one-stage model, learning is based on the difference between the imputed outcome for a decision $option_{i'}$ and the perceived probability that $option_{i'}$ would succeed or fail. The weights relating categories to evidence ($w_{i,j}$) are updated using an appropriately modified version of Equation (3) (i.e., $w_{i',j} = \eta \, \delta_{i'} \, category_j + w_{i',j}$) with δ_i' calculated as in Equation (4).

A difficulty arises for the two-stage model in specifying how decision makers learn categories since they usually do not receive direct feedback on which categories they should use. To address this difficulty, the two-stage model assumes that decision makers refine whatever may be their initial expectation of the categories that apply based on the categories' weighted performance in predicting the success or failure of given decision options. More formally, as

originally proposed by Rumelhart, Hinton, and Williams (1986), this assumption translates into the following calculation for δ_j, the value used to alter the weights ($w_{j,k}$) between the category and evidence layers:

$$\delta_j = \sum_{i=1}^{m} \delta_i \, w_{i,j} \tag{7}$$

with the rule for altering $w_{j,k}$ derived by again appropriately modifying Equation (3) (i.e., $w_{j,k} = \eta \, \delta_j \, evidence_k + w_{j,k}$).

Given that the two-stage model has been constructed as an extension of the one-stage model, it should lead to similar predictions as the one-stage model without transfer. The two models should differ in their predictions for transfer. Whether the difference is significant and substantive in a common dynamic decision task is a matter for empirical investigation.

3 Experiments

The goal of the experiments was to test the theoretical claims embodied in the one and two-stage learning models. For the sake of clarity, the method used in the human subjects experiment is first described and then related to the specific model instantiations. After that, the models' predictions for subject performance in the experimental task are compared and related to actual subject performance.

Method for Human Subjects

Fifty-four paid University of Michigan students participated in a four-session bargaining experiment over consecutive days. Subjects were paid $5 per session with a bonus of $40 if they completed all four sessions, making for a total payment of $60.

In the experiment, subjects dealt with opponents who accepted or rejected their most recent offer according to the sequence of offers that subjects had made up to that point. Learning to make the most favorable offer with the highest likelihood of success on any given speaker turn was possible if subjects tracked their last two offers and their opponents' responses to them.

During the first three sessions of the experiment, subjects dealt with opponents who responded one way according to particular sequences in their offers. In the fourth session, subjects dealt with opponents who responded to superficially different sequences of offers.

Subjects' understanding of what information to consider during bargaining was manipulated through three hint conditions applied between subjects: misleading, none, and good. For the first three days (sessions), this design led to three within-subjects measures of performance crossed by the three between-subjects hint conditions. On the fourth day, subjects' performance was measured in four consecutive intervals leading to a four-within by three-between design.

To help determine whether hint condition and experience influenced subject performance in the fourth session, 72 control subjects were run in a one-day experiment where they negotiated solely with opponents who responded as in the fourth session of the multi-session experiment. These control subjects were paid $15 for completing their one session experiment.

Detailed procedure. All sessions of the experiment took place using an Internet-based bargaining environment. Subjects came to a computer lab and used a web-browser to connect to a server that supposedly allowed them to interact with other subjects. Similar interactive web interfaces are becoming increasingly common in customer service and sales organizations.

At the start of the experiment subjects were informed that they would be playing the role of debt collectors against other "debtor" subjects who were behind on their credit card payments. These other subjects were in fact computer debtors. During each session, subjects made contact with 20 debtors whom they bargained with for twelve speaker turns lasting four seconds each, corresponding approximately to bargaining times observed in an actual field situation (Gibson, Fichman, & Plaut, 1996). Subjects' goal on each speaker turn was to get the debtor to agree to pay as much money in as short a time as possible. Even if debtors agreed to an offer, subjects had to continue to bargain because they might be able to achieve a more favorable agreement. Furthermore, debtors had proven poor in keeping past commitments so that, even if subjects were not able to achieve better terms, they had to keep proposing and re-proposing offers to reinforce

the debtors' commitment (Gibson et al., 1996; Sutton, 1991).

If subjects participated in the misleading or good hint conditions, they received the appropriate hint listed in Table 1 before each contact. Similarly to the effect information displays may have had in past experiments (e.g., Gibson, 2000), these hints were designed to influence the evidence (i.e., the pattern of offers and responses) subjects considered as they bargained. As indicated in the table, the "good" hint correctly told subjects that they needed to take into account their last two offers and the debtors' responses to them in order to make offers that had a high likelihood of success. The hint also provided subjects with two hypothetical sequences and the specific parts of those sequences they should consider. Neither the hint nor the example sequences gave an indication of what the correct offer after such sequences should be. They merely indicated the information that subjects should consider.

Condition	Details
good hint	When making an offer think of what your last two offers were and the debtor's response to them (accept or reject). For example, if you offered $100 in 8 days and the debtor rejected it and then $500 in 3 days and the debtor also rejected, that sequence may be significant for the offer you are about to make.
misleading hint	When choosing an offer, focus on whether you are on an speaker turn divisible by 2, 3, 5, or 7. For example, the debtor may be responding differently to the offer $500 in 3 days on speaker turns divisible by 7 than on speaker turns divisible by 3.

Table 1: Summary of Hint Conditions

The misleading hint and examples asked the subject to consider whether he or she was on a turn divisible by 2, 3, 5, or 7 and the debtor's last response. Thus, contrary to the good hint, this hint was designed to motivate subjects to consider a non-informative set of cues as evidence. Although, the hint was patently misleading, the empirical question remained as to its effect on subjects' behavior.

Subjects in the "none" hint condition received no hint about task structure, thus providing a comparison with the two other hint conditions.

Bargaining task.   Figure 2 shows the bargaining task interface. The task was simplified from field observations in a credit collections call center to make it more tractable for experimental work (Gibson & Fichman, 2001). As part of this process, it was designed to retain three features that relate to other functioning dynamic decision tasks. First, decision makers had to drive interactions under time pressure without fully controlling them (Brehmer, 1995; Kanfer & Ackerman, 1989; Klein et al., 1993; Sterman, 1994). Second, making a decision was essentially one of classifying the situation as to which offer to make (Gibson et al., 1996; Klein et al., 1993). Finally, debtors' responses showed patterns of sequential dependency making inference about the effectiveness of individual offers hard (Brehmer, 1995; Diehl & Sterman, 1995; Kleinmuntz, 1993; Sengupta & Abdel-Hamid, 1993; Sterman, 1989b).

On each speaker turn, subjects had four seconds to make their offer by clicking one of three options: L(ow) ($100 in 8 days); M(edium) ($300 in 5 days); or H(igh) ($900 in 2 days), and then clicking on a talk button. After that, subjects both saw on screen and heard through headphones the debtor's response of accept or reject to the offer.

This response was derived from the state transition diagrams (STDs) displayed in Figure 3. Debtor 1 (Figure 3(a)) was used for the first three sessions of training and Debtor 2 (Figure 3(b)) for the transfer session and one-session control subjects. STDs and similar approaches have been used to describe observations of sequential interaction in a number of functioning environments including customer service and marriage counseling (Gottman & Roy, 1990; Pentland & Reuter, 1994).

At the start of each contact, debtors were assumed to be in the state labeled $Reject_{(1)}$ in each of the STDs. As subjects made offers, the debtor changed to the next state along the labeled arrow that corresponded to the offer. After each such state change, the debtor responded according to the state's label of <u>accept</u> or <u>reject</u> 90% of the time and the opposite way the other 10% of the time.
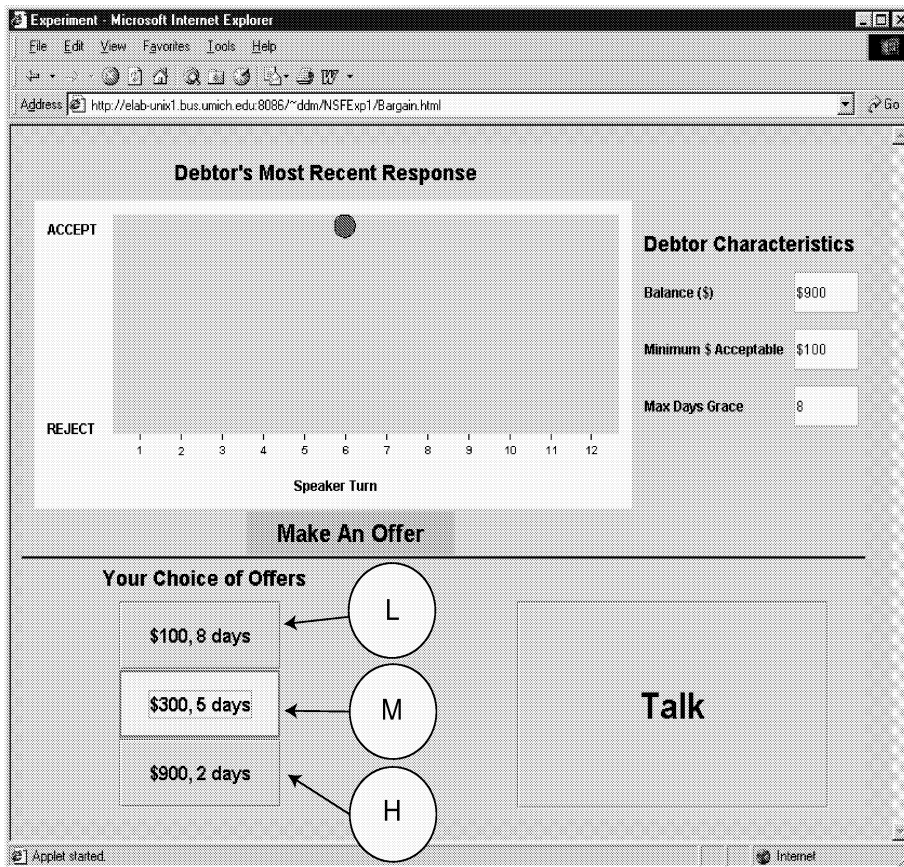
Figure 2: A single debtor contact. The figure shows a hypothetical subject's information display as she was selecting her seventh offer. For each contact, subjects were prompted to make twelve offers by clicking on one of the three buttons in the lower left-hand of the figure and then clicking the button labeled talk. The debtor responded by saying either accept or reject, and the red ball moved to the appropriate position. The circles labeled L(ow), M(edium), and H(igh) in the figure indicate the label that was used for each offer in the analysis and are presented as a convenience to the reader.
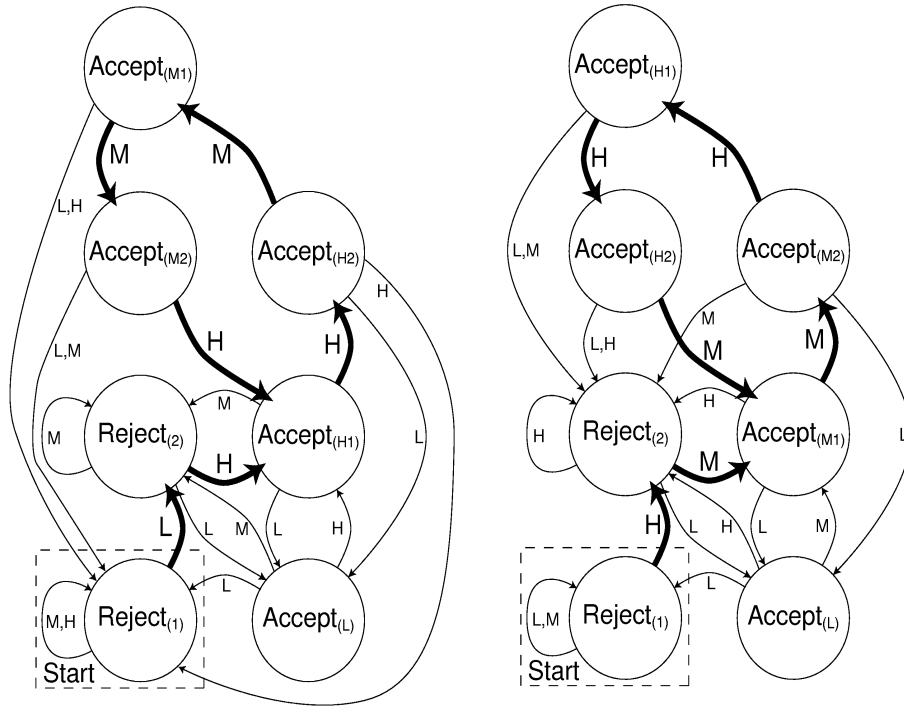
Path of best offers: ➝
Best-offer sequence: LH $\overline{\text{HMMH}}$

(a) Debtor 1

Path of best offers: ➝
Best-offer sequence: HM $\overline{\text{MHHM}}$

(b) Debtor 2

Evidence	State	Best Offer
$\text{Accept}_{(H)}\|\text{Accept}_{(M)}$	$\text{Accept}_{(H1)}$	H
$\text{Accept}_{(H)}\|\text{Accept}_{(H)}$	$\text{Accept}_{(H2)}$	M
$\text{Accept}_{(M)}\|\text{Accept}_{(H)}$	$\text{Accept}_{(M1)}$	M
$\text{Accept}_{(M)}\|\text{Accept}_{(M)}$	$\text{Accept}_{(M2)}$	H

(c) Analysis of repeating best-offer sequences.

Figure 3: State transition diagrams for Debtors 1 and 2 with a table of the common evidence and states in the repeating (overlined) portions of their best-offer sequences. See text for details.

For each state in each debtor, there was one offer that led to the highest immediate payoff, subjects' goal in the experiment. These best offers are indicated by a thicker arrow in each STD. After the first two offers for each debtor, subjects could enter into a repeating sequence of best-offers (overlined in the figures) that consisted entirely of high and medium offers.

The overall pattern of state transitions used in the debtor STDs was derived from foot-in-the-door (FID) and door-in-the-face (DIF) phenomena observed in functioning environments (Cialdini, 1984; Fern, Monroe, & Avila, 1986). In FID, bargainers get their opponents to accede to a relatively low request and then move them up to levels that they could not have obtained on one request alone. For instance, a college endowment drive might start out asking for small amounts of money and then, once those are agreed to, present requests for higher levels of commitment than could have been achieved on the first attempt. In DIF, bargainers make a burdensome request that will almost certainly be rejected and then get their targets to accept a request that they would not have accepted if made by itself without the first request. Real estate agents who start their customers out in neighborhoods at their maximum price limit are displaying one version of this strategy.

Debtor 1 in Figure 3(a) displayed elements of both the FID and DIF patterns. For this debtor, all of the subjects' first offers were rejected, suggesting a reticence frequently observed in the field. However, if the subject made a low offer on the first speaker turn, the debtor responded with a reject but would accept a high or low offer on the next turn (FID with an initial reject). If the subject made more than two high offers in a row after making a low offer on the first speaker turn, the third such offer would be rejected (delayed DIF). In this situation, only a retreat to medium or low would be accepted. Alternatively, after making a low offer on the first speaker turn, subjects could make another low offer and then build to high (FID).

Debtor 2 in Figure 3(b) also displayed elements of both the FID and DIF patterns but required a different offer sequence than Debtor 1. First, a high offer was required before the subject could start to get offers accepted (DIF). After the first rejected high offer, the subject had

to drop to a medium or low offer in order to have it accepted. After the first accepted medium offer, subjects could only make one other medium before they had to either raise it to high (FID) or drop to low (DIF). Similar levels of sequential dependency have proven difficult for subjects in other dynamic tasks (Berry & Broadbent, 1984; Gibson et al., 1997; Stanley et al., 1989).

Although the sequences of best-offers for each debtor STD may appear superficially quite different, the repeating portions of each sequence shared an identical underlying mapping of evidence to best offers, as shown in Figure 3(c). If subjects considered evidence consistent with the good hint (their most recent offer, the one before, and the debtors' responses), they could with practice determine the best offer with a high degree of certainty in the repeating portion of the sequence. For example, as illustrated in the first row of Figure 3(c), if either Debtor 1 or Debtor 2 had just accepted the subject's high offer after accepting a medium offer, the evidence indicated a situation in which the best offer was high. Subjects who learned evidence categories congruent with the rows in Figure 3(c) while bargaining with Debtor 1 could reuse them for better performance in bargaining with Debtor 2.

As further suggested by Figure 3(c), even the misleading hint might be of some use to subjects by recommending they consider evidence related to the point at which an offer occurred in the sequence. As apparent from the best offer sequences in Figures 3(a) and 3(b), a rote strategy of offer execution by position in the sequence could have been effective, although very difficult to attain by trial and error in practice (there were 3^{12} possible offer sequences in any given contact).

<u>Simulation Method</u>

The formulations for the one and two-stage models propose differing processes by which decision makers take decisions and learn from feedback but do not provide direct predictions that can be easily contrasted. To make such direct predictions, one and two-stage learning models were instantiated and run in a simulation experiment with Debtors 1 and 2.

Instantiating the learning models required: (1) determining the models' representation of the

offer options (high, medium, or low); (2) hypothesizing what evidence decision makers considered in making each offer; and (3), in the case of two-stage models, determining the number of categories. We address each of these in turn and then briefly examine the mechanics of how models made offers and learned in the simulation study.

The three offer possibilities were represented as separate action options. The principal evidence subjects used as they made offers was assumed to be of two types: (1) their last offer paired with the debtor's response; and (2) the hint if any. Subjects were further assumed to perceive offer-debtor-response pairs as a single discrete unit, a representation commonly assumed in cognitive game theory (Fudenberg & Levine, 1998, Chapter 4), repeated decision making (Dienes & Fahey, 1995), and skill automatization (Logan, 1988). For instance, if the subject made a high offer and the debtor rejected, the subject would encode that offer-response pair as high-offer-reject. In the simulation, a 1-of-n vector representation was used to capture this perceptual encoding. Since there were three possible offers and two possible responses, the vector had a length of six with 1 being placed in the position of the offer-response pair that had occurred and 0 in the other positions.[2] In the conditions where subjects received no hint, this 1-of-n representation of the last interaction was the only input the simulation received for each decision.

The hints were assumed to cause subjects to consider the additional evidence indicated. The good hint was represented by adding another 1-of-6 vector encoding for the offer-response pair that had occurred prior to the most recent pair. The misleading hint was instantiated by adding a four-place vector representing whether the turn was divisible by 2, 3, 5, or 7 and placing 1 in each position when true and 0 when not.

For the two-stage model, subjects were not assumed to have an a priori assumption about how many categories there might be. Therefore, in two-stage models, the number of non-exclusive categories that could be detected was varied between five and fifteen in increments of five. No

---

[2]Alternative representations are possible, in particular representing the last offer and debtor response separately. This and other alternative representations were tried without affecting the general pattern of simulation results reported presently.

effect was found on the overall pattern of results, but models capable of detecting 15 categories appeared to learn more reliably. Therefore, for two-stage learning, 15 categories were used.

At the start of each simulation experiment, the model weights were set to small random values close to 0. This assumption caused models to display individual learning characteristics since the different weights for each model represented different priors concerning which offer to select at the start of learning (Bishop, 1996). Therefore, eighteen models were run in each condition to estimate average performance.

Once instantiated, the models made offers and learned as follows. For each offer, the appropriate evidence vector was presented based on bargaining up to that point and the hint condition.[3] In two-stage models, the graded applicability of each evidence category was then computed and used to determine the relative support for each decision option. An option was then chosen randomly based on its relative support. For one-stage models, the offer process was foreshortened because the evidence vectors were used directly to determine relative support for the decision options without first being categorized.

Learning occurred as follows. After the offer was made, depending on which phase of the experiment the model was in, the Debtor 1 or Debtor 2 STD was used to produce a response. In the case of two stage models, this response was used to calculate a target for each offer option which was then used in the computations outlined earlier to change the weights between categories and offers. Then, as further described earlier, computations were performed to change the weights between the categories and and evidence, strengthening categories that performed well in predicting the success or failure of different decision options and weakening those that did not. Again, learning in one-stage models was simplified because only the weights between evidence vectors and decision options had to be altered. At this point, having learned by altering their weights based on the previous offer and the debtor's response to it, models were ready to consider new evidence and make another offer, repeating the process outlined above until the end

---

[3]At the start of each contact, no offers had been made, so the value for all offer-response pairs was set to 0.

of the contact.

<div align="center">Results and Discussion</div>

The models provided predictions for decision makers' best offer performance. Figure 4

summarizes the models' predictions and subjects' results for easy comparison.

The first three sessions.   Figure 4(a) shows the number of best offers made by one-stage

models, two-stage models and subjects over the first three sessions of the experiment. In each

session, perfect performance would have resulted in 240 best offers. Random guessing would

have led to 80 best offers (1/3 of total). On average, model and subject performance fell between

these two bounds.

Figure 4(c) presents the main effects of hint and practice on best offer performance for

models and subjects calculated using 1 *df* contrasts (Judd & McClelland, 1989). We focus first on

model predictions. Since both models assumed that subjects only considered the evidence

suggested by the hint, they both predicted that subjects receiving the good hint would outperform

those receiving the misleading hint. Next, since the misleading hint focused on evidence that

could help in determining position in the sequence relative to the none hint condition, both models

also predicted that subjects receiving any hint would outperform subjects receiving no hint.

Finally, both models predicted a positive linear trend in subjects' performance across sessions.

They differed in whether this trend would level between Sessions 2 and 3 (quadratic contrast),

with the one-stage model predicting a leveling of the trend and the two-stage model not.

As predicted by the models, subjects receiving the good hint significantly outperformed

those receiving the misleading hint. However, unlike models, subjects receiving any hint did not

perform significantly better than those in the none hint condition. Finally, as predicted by both

models, subjects showed a significant positive trend in performance over the three sessions.

Subjects failed to display a significant leveling of this trend as predicted by the one-stage model.

Looking more closely at the subject data, it appears that the misleading hint impaired
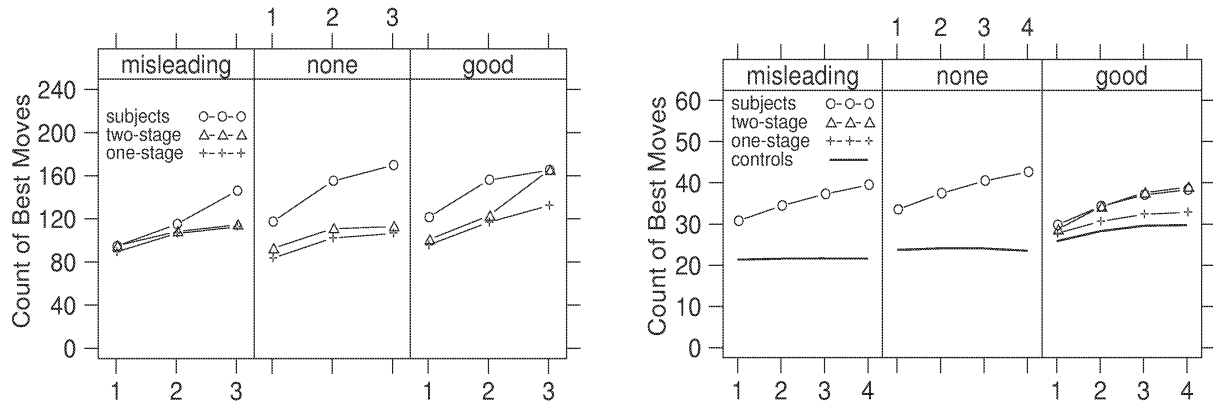
(a) Performance in Sessions 1–3.   (b) Performance in transfer (groups of 5 contacts).

	one-stage			two-stage			subjects		
	β	t_{51}	p	β	t_{51}	p	β	t_{51}	p
Hint									
Good vs. misleading	25.85	10.72	0.001	49.89	13.78	0.001	61.64	2.65	0.01
Any hint vs. none	28.08	11.64	0.001	29.94	8.27	0.001	−23.30	−1.50	ns
Trends									
Linear	27.70	29.84	0.001	34.78	28.25	0.001	49.26	9.42	0.001
Quadratic	10.41	7.39	0.001	1.22	0.66	ns	12.74	1.48	ns

(c) Predicted main effects of hint and experience in Sessions 1–3 (1 df contrasts).

	one-stage			two-stage		
	Δ	t_{17}	p	Δ	t_{17}	p
Misleading	46.94	2.31	0.03	37.50	1.84	0.08
None	149.50	5.83	0.001	125.33	4.89	0.001
Good	97.56	4.14	0.001	54.11	2.29	0.03

(d) Distance (Δ) of model from subject performance by hint in Sessions 1–3.

	one-stage			two-stage			subjects		
	Δ	t_{17}	p	Δ	t_{17}	p	Δ	t_{17}	p
Misleading	8.61	6.33	0.001	14.83	4.90	0.001	55.88	5.20	0.001
None	6.56	6.56	0.001	20.61	14.49	0.001	58.71	4.59	0.001
Good	45.61	11.43	0.001	55.56	7.74	0.001	26.17	2.63	0.02

(e) Difference (Δ) in performance between transfer and control performance.

	one-stage			two-stage		
	Δ	t_{17}	p	Δ	t_{17}	p
Misleading	18.00	1.68	ns	1.83	0.17	ns
None	30.17	2.35	0.03	14.00	1.09	ns
Good	16.00	1.61	ns	−0.20	−0.02	ns

(f) Distance (Δ) of good-hint models from subject performance by hint in transfer.

Figure 4: Training and transfer performance. See text for details.

subjects' performance rather than the good hint helping; subjects receiving the good hint did not significantly outperform those in the none hint condition over the three sessions. As shown in Figure 4(a), this impairment occurred primarily during the first two sessions. By the third session, the performance of subjects in the misleading hint condition was not significantly different from performance in the other two conditions.

Thus, the assumptions embedded in both models about how the hint would affect the evidence subjects considered proved incorrect on two counts. First, subjects in the none hint condition appear to have evolved within the first session to considering evidence of equivalent value to that suggested by the good hint. Second, although subjects' performance was impaired by the misleading hint, subjects with this hint climbed to performance by the third session that was statistically indistinguishable from that of subjects in the other two conditions, suggesting that they too evolved to considering evidence equivalent to the good hint.

These observations are reinforced by the analysis in Figure 4(d) of how well both models fit subject data. As apparent in the table, the greatest deviation for both models occurred in the none hint condition where they assumed that subjects only considered the last offer-response pair. Further, although both models were initially close to subject performance in the misleading hint, with overall two-stage model performance insignificantly different from that of subjects, the deviation of both models' performance from that of subjects grew as learning progressed. Only in the good hint condition did two-stage model performance converge toward that of subjects as learning progressed.

Transfer. Figure 4(b) shows subjects' transfer performance relative to control subjects as well as one and two-stage models' transfer performance in the good hint condition. Here, perfect performance would have resulted in 60 best offers per each group of five contacts shown in the figure and random performance in 20 best offers. Again, on average, models and subjects fell between these two bounds.

As indicated by the table in Figure 4(e), models in all hint conditions predicted that subjects

would outperform controls, and they did. Further, models, continuing with the original assumptions about how the hint influenced the evidence considered, predicted significant variation in transfer performance between hint conditions ($t_{51} = 4.31$, $p < 0.001$). However, subjects did not vary significantly in performance by hint condition, as might be expected given the insignificant difference in performance between subjects in the three hint conditions at the end of Session 3. As shown in Figure 4(f), the average transfer performance of subjects in any hint condition was closest to that of two-stage models using the good hint and insignificantly different from it

Two questions arise from this last observation. Are subjects in any hint condition and two-stage models with the good hint using similar or nearly similar processes in transfer? Are these processes distinguishable from those employed by the one-stage model with the good hint? To answer these questions, a detailed analysis was made of the performance of the best subject, two-stage model, and one one-stage model from the good hint condition. The subject's performance was equivalent to that of the five other highest performing subjects, one also in the good hint condition and two each in the other two hint conditions. Focusing on best performers provided an indication of the deficiencies that remained in models' best account of human behavior. Further, characteristics of good performance are of more interest to practice.

The best performing one and two stage models were able to achieve 50 best offers by the fourth set of five contacts in the transfer session. In this same period, the best subject was able to achieve 60 best offers, perfect performance. Over the full transfer session, the best subject made 215 best offers, the two-stage model 159 best offers, and the one-stage model 134 best offers.

In the analysis shown in Figures 5(a)–5(c), the STD for Debtor 2 has been depicted so that the weight of the arrows between the debtor's states indicates the relative frequency with which the subject, one-stage model, and two-stage model respectively made particular offers from that state during transfer. Examining the figures, an important difference between the human subject and both models is readily apparent for offers from the state labeled Reject$_{(1)}$. As shown in

Figures 3(a) and 3(b), Debtors 1 and 2 required different offers to transition from this start state

with the offer that worked for Debtor 1 (low) failing for Debtor 2 (required high). In Figure 5(a),

the slightly thicker arrow labeled <u>L, M</u> indicates that the subject took a few tries (8) to learn this

difference, but as shown by the thickness of this arrow in Figures 5(b) and 5(c), both models

required substantially more (66 and 40 respectively for two-stage and one-stage models).

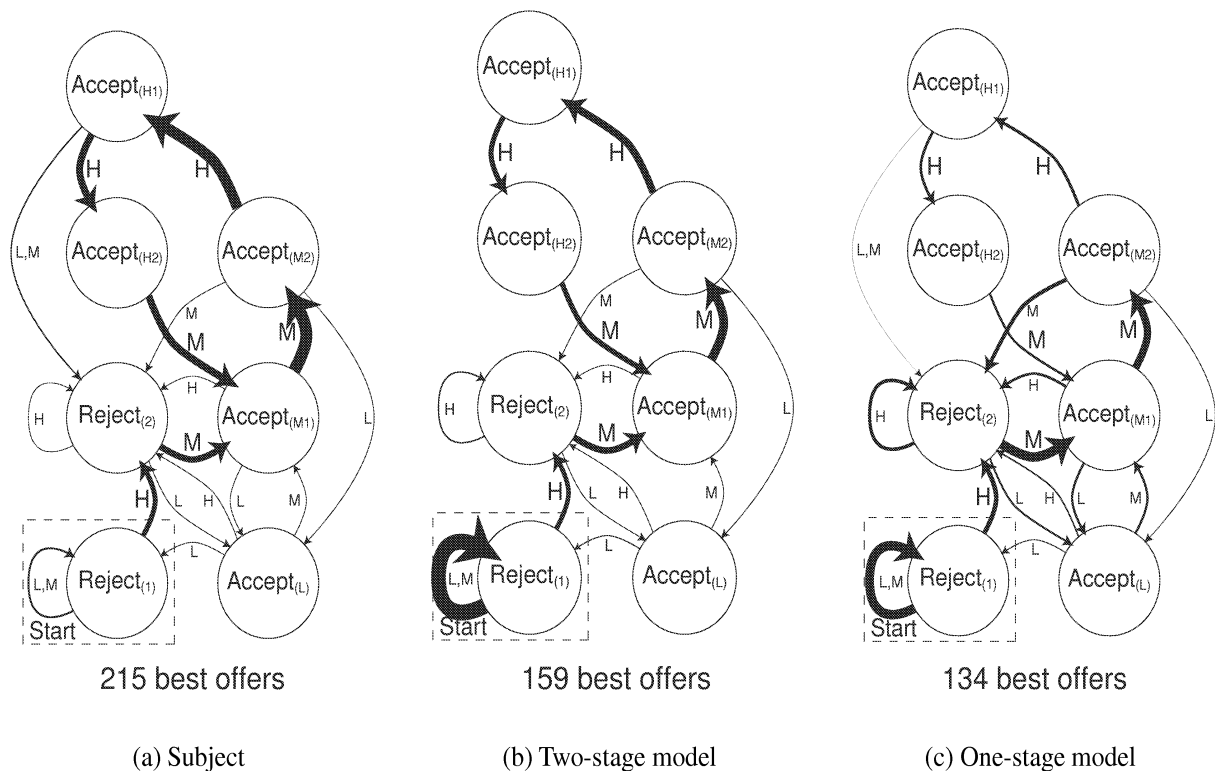| 215 best offers	159 best offers	134 best offers
(a) Subject	(b) Two-stage model	(c) One-stage model

Figure 5: Qualitative comparison of individual transfer.

However, after this initial difficulty, the two-stage model diverged from the one-stage model

to perform much more similarly to the subject on the repeating portion of the best-offers

sequence. While both the two-stage model and the subject made occasional errors as indicated by

the thin arrows that depart from the repeating path of best offers, the vast majority of their offers

in this part of the STD were best offers. Such was not the case for the one-stage model. First, it

took more tries to relearn to transition from the state labeled $Reject_{(2)}$ than either the subject or

two-stage model (14 vs. 2 and 5, respectively). Next, it showed a higher reject rate from the state

labeled Accept$_{(M1)}$ (10 vs. 2 and 3) because it attempted to offer high in that state, the offer that would have worked from the same position in the sequence in Debtor 1. This trend continued in Accept$_{(M2)}$ where the one-stage model made more medium offers (16 vs. 1 and 3), the best offer from this position in the sequence in Debtor 1, than high offers (12 vs. 49 and 32), the best offer from this state given the evidence category.

Thus, while both models had more initial difficulty than the subject in relearning the offer it took to get beyond Reject$_{(1)}$, the two-stage model subsequently tracked the subject's performance much more closely. The two-stage model's performance in the rest of the STD was due to the fact that it had learned to categorize offer patterns so that it recognized the repeating portion of the best offer sequence, identical across Debtors 1 and 2. The one-stage model appears to have had to relearn this path because it more frequently made the offer that would have applied in the same sequential position (but not the same category as determined by the pattern of offers and responses) in the Debtor 1 STD.

The question remains as to whether the subject learned categories similarly to the two-stage model or simply relearned the STD for Debtor 2, just more quickly than the one-stage model. It seems most likely that the subject recognized the categories from patterns of offers and responses. In the one case, Reject$_1$, where the subject had to relearn what offer to make, he made substantially more errors than anywhere else in the STD. Had the subject relearned the full STD, this higher error rate should have persisted throughout.

<u>Summary.</u> Both one-stage and two-stage models predicted elements of decision makers' behavior in the first three sessions, with the two-stage model providing better fits to subject data. As predicted by both models, decision makers receiving good hints outperformed those receiving misleading hints, and decision makers showed a positive linear trend in performance, though not the leveling off in performance predicted by the one-stage model. However, both models were proven largely incorrect in their initial assumptions concerning how the hint would affect the evidence decision makers considered. By the end of the first three sessions, decision makers

converged to an equivalent high-level of performance that suggested they were considering the evidence highlighted by the good hint.

This pattern persisted in the transfer task where average two-stage model performance with the good hint was not significantly different from subject performance in all conditions. A detailed analysis of the best performing subject, two-stage model, and one-stage model suggested that high-performing subjects were able to transfer because they had learned to categorize patterns of offers and responses that were identical between the first three sessions and transfer, according well with the theoretical explanation provided by the two-stage model with good hint. However, subjects' behavior was at variance with the best-performing one-stage model that had to essentially relearn the task. Thus, in this task, high-performing decision makers' behavior appears to provide support for the central premise of the two-stage model: <u>decision makers in dynamic tasks learn to categorize patterns of behavior, are able to recognize these categories in novel situations, and use these categories to cue behaviors that were successful in the past</u>.

<div align="center">4   General Discussion and Conclusion</div>

An important feature of the results just reported is the degree of transfer exhibited by human subjects. Both the one and two-stage models predicted aspects of subjects' performance in the three sessions of the experiment prior to transfer. Only the two-stage model proposed a process, based on evidence categorization, to account for how subjects could transfer knowledge they had learned in the first three sessions to the transfer task. Subjects' behavior in transfer best fit the two-stage model's account when the proper assumption was made about the evidence they were considering. Detailed comparison between both models and a high performing subject indicated that the two-stage model's categorizations accorded well with those used by the subject. Unlike the one-stage model, the high performing subject did not have to go through extensive relearning when, in a novel situation, he re-encountered categories of evidence he had experienced in the first three sessions.

<center>Limitations</center>

These findings have important implications for research and practice. However, there are limitations in the models, task, and procedures used in this study that need to be addressed before discussing them.

Models. As apparent in the results reported above, the models faced three major limitations in accounting for subject performance. First, both models assumed that the evidence people considered was fixed and determined by the hint. Although this assumption provided a clear distinction between the different information scenarios, subjects in both the none and misleading hint conditions appeared to adapt in the information they considered such that their aggregate performance fit best with the two-stage model using the good hint by the end of the first three sessions and during transfer. One possibility to account for this pattern of results is that the supposed interaction with another subject cued subjects to consider sequences of offers and responses, since this is a natural feature of such interactions, even when subjects know the interaction is faked (Berry & Broadbent, 1984). Under this scenario, the misleading hint would have added an alternative set of hypotheses that subjects would have had to weigh against others with the gradual accumulation of evidence, thereby slowing learning (e.g., Gibson & Plaut, 1995). However, this view runs counter to the other frequent result that decision makers in dynamic tasks fail to consider sequential dependencies (e.g., Sterman, 1989b). Thus, the information decision makers consider and how it evolves as they perform in dynamic tasks remains a topic for future research.

Second, in comparing individual transfer performance, both models took substantially longer than the subject to relearn the new high offer required at the start of bargaining. A common modeling solution to correct this deficiency is to supply a "forgetting" parameter (Erev & Roth, 1998) that allows the model to show a level of adaptation away from previously learned responses more in line with that of decision makers. However, introducing such a global parameter can have

a blurring effect on which of the model's claims are due to transfer of prior knowledge and which due to rapid relearning and so was not used here.

Finally, as apparent in their construction, both models start from the assumption that learning in dynamic tasks is essentially a process of accumulating weights for categories and evidence in memory. The models further assume that this process is uniform and systematic across individuals. As such, the models only provide a limited account of between-subjects variation in performance that is related to different prior beliefs and accidents of their experience in the task. Other models with such a parsimonious focus have been argued to account for a large part of expert performance in games (e.g., chess, Gobet & Simon, 2000), and the reliance on recall has generally been cited as a feature of performance in functioning dynamic environments (Klein et al., 1993). However, memory-focused accounts such as the one proposed here leave aside factors such as motivation and problem solving capacities that have proven important in the early stages of learning in other studies of functioning dynamic environments (Kanfer & Ackerman, 1989). Incorporation of these other factors into the models examined here is left to future work.

Task and procedures.   The task and procedures also have three principle limitations. First, the task is simple relative to some tasks studied elsewhere (e.g., Diehl & Sterman, 1995; Sterman, 1989b) and relative to the field environment from which it was derived. These limitations were imposed to make the task more tractable in the time available to subjects, but may also suggest limitations on the generalizeability of the findings. In particular, the behavior of both debtors was relatively stable and decomposable. Diehl and Sterman (1995) investigated a market planning task in which the environment was much more subject to change. Although they reported some improvement with practice, their subjects did not show the degree of learning exhibited here. Thus, an important potential limitation on the findings reported here is that they require a relatively stable environment to be realized.

As for decomposability, both debtors were constructed such that their internal states were easily recognized using two offer-response pairs. Although many functioning environments

appear to have similarly short sequential dependencies (e.g., marriage counseling and customer service, Gottman & Roy, 1990; Pentland & Reuter, 1994), the same cannot be said of all dynamic decision environments. For instance, Sterman's (1989b) beer distribution task starts with short delays between order and delivery that then increase due to decision makers' ill-considered efforts to meet a perturbation in demand. Sterman's decision makers appeared not to be able to adapt to the longer dependency. How decision makers adapt to longer dependencies is an important topic for future research. The results reported here are encouraging because they suggest that decision makers are able to adapt to shorter dependencies that may be common to many environments.

Second, debtors' responses lacked any emotional content, although other work suggests that emotions are frequently displayed in this type of interaction and may influence outcomes (e.g., Gibson & Fichman, 2001; Sutton, 1991). However, a primary activity of workers involved in such interactions is to engage in exercises that limit emotion's role. To the extent that emotion can be successfully compartmentalized, the work reported here should remain applicable.

Finally, although the misleading hint reduced performance as predicted, it lacked face validity. Further, the good hint did not lead to significantly better performance over the none hint condition. Thus, although the results reported here strongly suggest that subjects performed well in transfer because they were able to make effective use of categories that they had learned earlier in the task, the manipulations presented have yet to be fully proven as a practical means for manipulating this process. Future work should continue to focus on effective means for influencing decision maker's categorization processes.

Implications

In spite of the limitations just listed, the work described here presents an important new perspective for promoting learning in dynamic decision environments. Previous reports have largely enumerated the many ways decision makers fail to learn in these environments or suggested severe limitations on their learning. Initially, Sterman and his collaborators reported on

the sub-par performance of executives and graduate students deemed to have useful practical and theoretical knowledge of dynamic environments (e.g., Sterman, 1989b). Even when subjects were explicitly instructed to make careful use of dynamic systems coursework and analysis, they failed to perform at a level that could be achieved with informed heuristics (Diehl & Sterman, 1995). Subjects were observed to throw down their pencils and paper and trust to intuition early on in the task.

In accord with this observation, Roth and Erev (1995) proposed that decision makers learn from success and failure conditioned on the specific evidence available at the time of the decision. While this account has provided good fits to data in a number of domains including skill automatization (e.g., Logan, 1988) and dynamic decision making (e.g., Dienes & Fahey, 1995), it suggests that decision makers can only apply their skills to exactly what they have seen before. As shown by the one-stage model in this study, this perspective implies that application of skilled performance to novel situations is largely a matter of relearning.

By counter, the work reported here presents evidence and a theoretical perspective embodied in the two-stage model that performance by skilled decision makers in novel situations is not a matter of relearning but rather recognition of evidence categories they have encountered before. As such, this is one of the first times that observation of laboratory behavior in dynamic tasks has made contact with theories of expert performance (e.g., Gobet & Simon, 2000; Simon & Gobet, 2000) and descriptions of functioning environments (Klein et al., 1993) that have long held this view. This study extends these prior efforts by explicitly demonstrating the link between task experience and later skill transfer in a fairly short amount of time. A direct implication of this link is that management interventions to improve categorization performance can be a major lever in improving decision makers' performance in the shorter term.

In this regard, at least three managerial interventions can be derived from this study. First, managers can attempt to organize environments so that they traverse periods of stability, a feature of the experimental task previously noted as a limitation. As observed in the functioning

environment from which the task was developed, managers grouped debtors by level of delinquency, and within these levels, collectors further grouped debtors who had specific types of financial problems or particular levels of accumulated debt. In this study, dealing with such stable characteristics allowed subjects time to learn evidence categories that they were then able to effectively reuse in a novel situation. The strong suggestion is that managers might be able to increase knowledge transfer between different types of interactions by encouraging this already existing tendency to inject stability into the work environment. On the flip side, managers need to consider the possibility of <u>unwanted transfer</u> that such practices may foster when evidence categories have different implications in different types of interaction.

Another possible intervention is the use of hints. In the task used in this study, decision makers appeared able to use the pattern of interaction highlighted by one of the hints to form evidence categories that they were then able to recognize in a novel situation and use effectively. Functioning environments such as the one from which this task was developed make extensive use of hints in the form of coaching tips from managers. In decision environments where much of the interaction is verbal and cannot be captured graphically, verbal hints and coaching are one of the few interventions available to help decision makers on the spot. This study's results suggest that the most effective hints may be oriented toward highlighting critical patterns of interaction from which the decision maker may draw his or her own inferences. This approach will develop skills in the decision maker which he or she is more likely to trust and use than "black box" decision aids that just tell the decision maker what to do (Little, 1970).

Finally, a frequent problem for learning in many dynamic tasks is that long periods may pass in between times when critical skill sets are needed suggesting that experience in the functioning environment alone may be insufficient to acquire them (e.g., battlefield command and air traffic control, Klein et al., 1993). Past research suggests two ways of addressing this problem. First, managers might try training decision makers in the theory of how dynamic environments function and then expect them to apply the theory in rarely occurring novel situations. However, it is in

precisely this application that previous work has shown decision makers to be the most deficient (e.g., Diehl & Sterman, 1995). Second, as suggested by the theoretical perspective instantiated in the one-stage model, managers might try constructing simulated training environments that exactly duplicate the situations that will be faced. However, anticipating every possible permutation that needed to be addressed would be extremely difficult. The theoretical perspective adopted in this study suggests that it is more profitable to anticipate and categorize the types of novel situations decision makers will face and develop simulated environments that train them to recognize these situation categories through experience. For this perspective to bear fruit, more research needs to be invested in the design and validation of such environments. Such research should be facilitated, up to and including final implementation, by the ongoing computerization of many dynamic decision environments.

Conclusion

Examinations of dynamic decision behavior in the laboratory have tended to indicate strong deficits in learning that do not always accord well with field observations. In particular, decision makers in functioning dynamic environments such as firefighting seem able to recognize categories of evidence in novel situations that they can then use to construct effective solutions (Klein et al., 1993). This study has presented two contrasting theoretical perspectives on decision makers' categorization ability in the one-stage and two-stage models and tested them against decision behavior in a task derived from a field setting. The one-stage model hypothesized that decision makers would essentially just learn idiosyncratic behaviors specific to the decision context they were exposed to. In contrast, the two-stage model hypothesized that decision makers would learn evidence categories that they could recognize in novel situations.

Both models displayed deficits in their account of subject behavior. However, the two-stage model's evidence categorization mechanism appears to have captured an important element of decision makers' behavior when they encountered a novel bargaining situation. Thus, the

theoretical perspective instantiated in the two-stage model provides the first elements of a framework for guiding the design of decision environments to promote effective learning. Further work is needed with this perspective to suggest and test ways it can influence managerial practice beyond the interventions proposed here.

References

Berry, D. C., & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. Quarterly Journal of Experimental Psychology, 36A, 209–231.

Bishop, C. M. (1996). Neural networks for pattern recognition. New York: Oxford University Press.

Brehmer, B. (1995). Feedback delays in complex dynamic decision tasks. In P. Frensch, & J. Funke (Eds.), Complex problem solving: The European perspective. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cialdini, R. B. (1984). Influence: The psychology of persuasion. New York: William Morrow.

Diehl, E., & Sterman, J. D. (1995). Effects of feedback complexity on dynamic decision making. Organizational Behavior and Human Decision Processes, 62(2), 198–215.

Dienes, Z., & Fahey, R. (1995). Role of specific instances in controlling a dynamic system. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(4), 1–15.

Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. American Economic Review, 88(4), 848–881.

Fern, E. F., Monroe, K. B., & Avila, R. A. (1986). Effectiveness of mulitple request strategies: A synthesis of research results. Journal of Marketing Research, 23, 144–152.

Fudenberg, D., & Levine, D. K. (1998). The theory of learning in games. Cambridge: The MIT Press.

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., & Fichman, M. (2001). Emotions as information in bargaining: What happens when the credit collector calls? Unpublished Manuscript, (in revision).

Gibson, F. P., Fichman, M., & Plaut, D. C. (1996). Learning in dynamic decision tasks—Credit collections at Citicorp (Technical report). Citicorp Behavioral Sciences Research Council.

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

Gibson, F. P., & Plaut, D. C. (1995). A connectionist formulation of learning in dynamic decision-making tasks. Proceedings of the 17th Annual Conference of the Cognitive Science Society (pp. 512–517). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gobet, F., & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. Cognitive Science, 24(4), 651–682.

Gottman, J. M., & Roy, A. K. (1990). Sequential analysis: A guide for behavioral researchers. Cambridge: Cambridge University Press.

Jarvenpaa, S. L. (1989). The effect of task demands and graphical format on information processing strategies. ManagementScience, 35(3), 285–303.

Joslyn, S., & Hunt, E. (1998). Evaluating individual differences in response to time-pressure situations. Journal of Experimental Psychology: Applied, 4(1), 16–43.

Judd, C. M., & McClelland, G. H. (1989). Data analysis: A model comparison approach. San Diego, CA: Harcourt Brace Jovanovich.

Kanfer, R., & Ackerman, P. (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., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.). (1993). Decision making in action: Models and methods. Norwood, NJ: Ablex Publishing Corporation.

Kleinmuntz, D. N. (1993). Information processing and misperceptions of the implications of feedback in dynamic decision making. System Dynamics Review, 9(3), 223–237.

Little, J. D. C. (1970). Models and managers: The concept of a decision calculus. Management Science, 16, B466–B485.

Logan, G. D. (1988). Toward an instance theory of automatization. Psychological Review, 95(4), 492–527.

McClelland, J. L. (2001). The mit encyclopedia of the cognitive sciences, Chap. Cognitive Modeling, Connectionist. Cambridge, MA: MIT Press. available on-line at http://cognet.mit.edu/MITECS/Entry/mcclelland.html.

Paich, M., & Sterman, J. D. (1993). Boom, bust, and failures to learn in experimental markets. Management Science, 39(12), 1439–1458.

Pentland, B. T., & Reuter, H. H. (1994). Organizational routines as grammars of action. Administrative Science Quarterly, 39(3), 484–510.

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior, 8, 164–212.

Rumelhart, D. E., Durbin, R., Golden, R., & Chauvin, Y. (1995). Backpropagation: The basic

theory. In Y. Chauvin, & D. E. Rumelhart (Eds.), Back-propagation: Theory, architectures, and applications. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(9), 533–536.

Sengupta, K., & Abdel-Hamid, T. K. (1993). Alternative conceptions of feedback in dynamic decision environments: An experimental investigation. Management Science, 39(4), 411–428.

Simon, H. A., & Gobet, F. (2000). Expertise effects in memory recall. Psychological Review, 107(3), 593–600.

Stanley, W. B., Mathews, R. C., Buss, R. R., & Kotler-Cope, S. (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. (1989a). Misperceptions of feedback in dynamic decision making. Organizational Behavior and Human Decision Processes, 43, 301–335.

Sterman, J. D. (1989b). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. Management Science, 35(3), 321–339.

Sterman, J. D. (1994). Learning in and about complex systems. System Dynamics Review, 10(2–3), 291–330.

Stone, G. O. (1986). An analysis of the delta rule and the learning of statistical associations. In Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations (pp. 444–449). Cambridge, MA: MIT Press.

Sutton, R. I. (1991). Maintaining norms about expressed emotions: The case of bill collectors. Administrative Science Quarterly, 36, 245–268.