ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
Vol. 71, No. 1, July, pp. 1–35, 1997
ARTICLE NO. OB972712

Learning in Dynamic Decision Tasks: Computational Model and Empirical Evidence

Faison P. Gibson

University of Michigan Business School, Ann Arbor

Mark Fichman

Graduate School of Industrial Administration, Carnegie Mellon University

and

David C. Plaut

Departments of Psychology and Computer Science, Carnegie Mellon University

Dynamic decision tasks include important activities such as stock trading, air traffic control, and managing continuous production processes. In these tasks, decision makers may use outcome feedback to learn to improve their performance "on-line" as they participate in the task. We have developed a computational formulation to model this learning. Our formulation assumes that decision makers acquire two types of knowledge: (1) How their actions affect outcomes; (2) Which actions to take to achieve desired outcomes. Our formulation further assumes that fundamental aspects of the acquisition of these two types of knowledge can be captured by two parallel distributed processing (neural network) models placed in series. To test our formulation, we instantiate it to learn the Sugar Production Factory (Stanley, Mathews, Buss, & Kotler-Cope, *Quart. J. Exp. Psychol.*, 1989) and then apply its predictions to a human subjects experiment. Our formulation provides a good account of human decision makers' performance during training and two tests of subsequent ability to generalize: (1) answering questions about which actions to take to achieve goals that were not encountered in training; and (2) a new round of performance in the task using one of these new goals. Our formulation provides a less complete account of

Address correspondence and reprint requests to Faison P. Gibson, University of Michigan Business School, 701 Tappan Street, Ann Arbor, MI 48109-1234. E-mail: fpgibson@umich.edu.

0749-5978/97 $25.00

decision makers' ability after training to predict how prespecified actions affect the factory's performance. Overall, our formulation represents an important step toward a process theory of how decision makers learn on-line from outcome feedback in dynamic decision tasks.

1. INTRODUCTION

Dynamic decision tasks include important activities such as stock trading, air traffic control, and managing continuous production processes. These *dynamic* decision tasks can be distinguished from one-time decision tasks, such as buying a house, by the presence of four elements: (1) The tasks require a series of decisions rather than one isolated decision; (2) The decisions are interdependent; (3) The environment changes both autonomously and as a result of decision makers' actions; (4) Decisions are goal-directed and made under time pressure, thereby reducing the decision-maker's opportunities to consider and explore options (Brehmer, 1990, 1992, 1995; Edwards, 1962; Rapoport, 1975).

These four properties frequently render dynamic decision tasks analytically intractable from a practical standpoint (Klein, Orasanu, Calderwood, & Zsambok, 1993). However, an important feature of many dynamic decision tasks is that they provide decision makers with feedback on the outcomes of their actions. Decision makers may use this outcome feedback to learn to improve their performance "on-line" as they participate in the task even without an analytic solution (Hogarth, 1981). This said, a consensus has emerged among researchers that decision makers in dynamic tasks fail to adjust their behavior on-line in response to feedback in a way that fully takes into account the task structure (Brehmer, 1995; Sterman, 1994; Kleinmuntz, 1993). To inform efforts to improve decision makers' learning in these tasks, Brehmer (1995) has called for a theory that addresses how decision makers process outcome feedback and learn from it on-line.

As a first step toward such a theory, we have developed a computational formulation that builds on previous theoretical work in dynamic decision making by Brehmer (1990, 1992, 1995) and motor learning by Jordan and Rumelhart (1992; Jordan, 1992, 1996; Jordan, Flash, & Arnon, 1994; Wolpert, Ghahramani, & Jordan, 1995). Our formulation proposes two central assumptions (Gibson & Plaut, 1995). First, decision makers use outcome feedback to form two interdependent, internal submodels of the task as they participate in it. These two submodels represent knowledge about: (1) how the decision maker's actions affect outcomes, and (2) which actions to take to achieve desired outcomes. Our formulation's second major assumption is that the acquisition of these two types of knowledge from outcome feedback can be simulated by on-line learning in parallel distributed processing (PDP) or neural network models.

The advantage of constructing a computational formulation is that we can use it to instantiate the assumptions we have just presented to generate testable predictions about human performance in different task manipulations and settings. We do this using the Sugar Production Factory (SPF), a simple dynamic decision task in which subjects, using outcome feedback alone, learn to manipulate an input (typically workforce) to achieve target levels of sugar production. The task has been widely used to investigate hypotheses about how decision makers learn on-line in dynamic decision environments (Berry, 1991; Berry & Broadbent, 1984, 1987, 1988; Berry & Dienes, 1993; Buchner, Funke, & Berry, 1995; Dienes, 1990; Dienes & Fahey, 1994, 1995; Gibson, 1996; Gibson & Plaut, 1995; Hayes & Broadbent, 1988; Marescaux, Luc, & Karnas, 1989; McGeorge & Burton, 1989; Sanderson, 1990; Stanley, Mathews, Buss, & Kotler-Cope, 1989; Squire & Frambach, 1990).

In studies using the SPF, subjects reliably learn to improve their performance using outcome feedback but show much less reliable improvement in subsequent measures of their ability to generalize to situations not encountered during training. Measures of generalization ability include subjects' ability to predict the factory's response to novel inputs and to specify inputs that will drive the factory to the production goal in situations they did not encounter during training. The poor ability to generalize based on learning from outcome feedback in the SPF is characteristic of others' observations that decision makers in dynamic environments fail to fully take into account task structure (e.g., Diehl & Sterman, 1995; Brehmer, 1995; Sterman, 1989).

Our formulation suggests three properties of learning from outcome feedback that may explain poor generalization in dynamic decision tasks like the SPF:

- Learning is approximate. Decision makers become more accurate but do not eliminate error in specifying actions to achieve goals.
- Learning is most applicable locally. Decision makers are the most accurate in specifying actions to achieve goals near the area of their greatest experience.
- Because of this approximate, local learning, transfer of knowledge to achieving a new goal is graded by its proximity to the training goals.

In the first study we report, we perform a computational simulation in which our formulation learns to control the SPF by using outcome feedback to adjust its decisions on-line. We use the results of the simulation to predict the course of subjects' performance during training as well as three measures of their ability to generalize from their experience after training: (1) *control:* how subjects will differentially respond to questions where they are asked to provide inputs to achieve new production goals; (2) *prediction:* how subjects will differentially respond to questions where they are asked to predict the factory's response to specific inputs given different levels of current production; and (3) *transfer:* how subjects will perform in the SPF when they are asked to achieve goals they did not experience during training. In the second study reported below, we test these predictions using a human subjects version of the task.

2. THE SUGAR PRODUCTION FACTORY

Subjects in the SPF manipulate an input to a hypothetical sugar factory to attempt to achieve a particular production goal. At every time step t, subjects must indicate the input (measured in hundreds) for time $t + 1$ and are usually limited to 12 discrete values ranging from 100 to 1200. Similarly, the output of the factory is bounded between 1000 and 12000 (generally measured as tons of sugar) in discrete steps, and is governed by the following equation (which is unknown to subjects):

$$P_{t+1} = 2W_{t+1} - P_t + \epsilon, \tag{1}$$

where P_{t+1} represents the new production at time $t + 1$ (in thousands), W_{t+1} is the input specified at $t + 1$ (in hundreds), and ϵ is a uniformly distributed random error term, taking on the values of -1, 0, or 1. Over a series of such trials within a training set, subjects repeatedly specify a new input and observe the resulting production level, attempting to achieve the prespecified production goal.

Note that subjects with knowledge of Equation 1 are always able to reach within 1000 of the production goal by rearranging its terms and substituting as follows to directly compute the input:

$$W_{t+1} = \frac{P^*_{t+1} + P_t}{2}, \tag{2}$$

where P^*_{t+1} is fixed at the production goal the subject wants to achieve. The random element, ϵ, does not figure in Eq. (2) because, by definition, it cannot be controlled.

The SPF has dynamic elements that are challenging to subjects who are attempting to learn it using outcome feedback alone (Berry & Broadbent, 1984, 1988; Stanley *et al.*, 1989). In particular, due to the lag term P_t, two separate, interdependent inputs are required at times t and $t + 1$ to reach steady-state production. Furthermore, there are two elements of autonomous change in the system that the subjects do not directly control by their actions. First, because the lag term P_t influences P_{t+1}, subjects must learn to condition their actions W_{t+1} on P_t to reach a given goal P_{t+1}. Otherwise, P_{t+1} oscillates. Second, the random element forces subjects to adapt to unanticipated changes in the SPF's state. If subjects are allowed to set their input in increments of 50, the random element also bounds the expected percentage of trials within one thousand of goal performance to between 11% (for randomly selected input values; Berry & Broadbent, 1984) and 100% (for a perfect model of the system; Stanley *et al.*, 1989).

Notice from this description that the SPF contains all of the elements of more general dynamic decision-making tasks defined by Brehmer (1990, 1992, 1995), with the exception of time pressure. One effect of time pressure is to limit decision makers' ability to consider and explore options, forcing them to react immediately to task contexts with the options available (Klein *et al.*, 1993). This effect of time pressure is captured explicitly in our computational formulation of learning described in Section 3. In the human subjects study, we add a form of time pressure to the SPF to drive subjects to react immediately to environmental contexts.

2.1. Previous Empirical Results

An important pattern of results that recurs in studies using the Sugar Production Factory is that while subjects improve their ability to reach the production goal with moderate experience using outcome feedback alone, other measures of their knowledge about the task improve much less reliably (Gibson, 1996). These measures include subjects': (1) correctness in predicting the SPF's response to specific input values in given levels of current production (Berry & Broadbent, 1984, 1987, 1988; Berry & Dienes, 1993; Buchner *et al.*, 1995), (2) performance in providing inputs to drive the factory to a production goal given the current production (Dienes & Fahey, 1995; Marescaux *et al.,* 1989), and (3) ability to generate useful heuristics for the task (McGeorge & Burton, 1989; Stanley *et al.,* 1989).

In their original work with the SPF, Berry and Broadbent (1984, 1988) found that while subjects improved in their ability to reach the production goal with practice, their subsequent ability to correctly answer prediction questions about the task was no better than that of subjects who had no task experience. In the prediction questions, subjects were presented on paper with three periods of past production and workforce values as well as the currently set workforce and asked to predict the next period's production. Answers were counted as correct if they were within 1000 of the value computed substituting the relevant data into Eq. (1) and computing the result before application of the random element.

However, using a somewhat more complicated variant of the task than that presented by Eq. (1), Berry and Broadbent (1987, p. 13) found that subjects learned with experience to better predict the direction of change for relationships that did not accord with their prior experience when the prediction questions were presented in the same modality that subjects had used to perform the task. This result suggests that, within the context of a given task, subjects do develop approximate knowledge about task responses to their inputs based on experience with the task and accords well with Funke's (1992) findings in other, similar dynamic decision tasks. Further contrasting with the initial results on prediction, Buchner *et al.* (1995) discovered that subjects had a higher likelihood of predicting correctly, using Berry and Broadbent's (1984) measure, if they had seen the question situation during training. Otherwise, these subjects had a probability of answering correctly similar to that of subjects who had received no training.

Together, Buchner *et al.*'s (1995) and Berry and Broadbent's (1987) results suggest that subjects may develop their ability to predict system responses most accurately, but still not perfectly, in the region of their greatest experience. Prediction ability degrades quickly with distance from the training region to the point where subjects with training appear no better than subjects without training using Berry and Broadbent's (1984) measure of correct performance.

Marescaux *et al.* (1989) and Dienes and Fahey (1995) tested the *concordance* of the workforce values subjects provided to questions after 80 trials of training on the SPF with the workforce values they had specified during training.[1] As in the prediction questions we have just described, each of these questions presented subjects with the current level of sugar production, the current workforce level, and the past three production values. Subjects had to specify the workforce that would bring production to goal. Dienes and Fahey (1995) found that in questions where current workforce and current production were both within one level of a situation encountered during training and where the workforce given in the training situation had led to hitting the production goal, subjects tended to provide the workforce they supplied during training in response to the question (*positive concordance*). Otherwise, subjects tended not to provide the workforce they supplied during training in response to the question (*negative concordance*). These results suggest that subjects: (1) are able to improve their accuracy in specifying workforce using outcome feedback without fully eliminating error and (2) develop a limited ability to generalize based on their experience learning from outcome feedback. This latter ability is reduced with increased distance from the training region. The limited ability to generalize mirrors Sterman's (1989) anecdotal finding that subjects learning in a much more complex dynamic decision task were able to improve their performance with experience but had very limited ability to transfer any knowledge they had gained to new situations with the same task.

Stanley *et al.* (1989) asked subjects in their *original learners* condition to write rules for other, naïve yoked subjects after each set of 10 training trials. This process was repeated for 60 sets of 10 trials. The rules of 11 subjects who achieved a performance criterion were systematically tested on naïve subjects. Stanley et al. found that only the rules of 3 of the 11 rule-writing subjects had a positive impact on naïve subjects' performance. Additionally, for the subjects who wrote rules, Stanley *et al.* determined breakpoints after which these subjects' performance increased at a faster rate than before. However, subjects' rules after the breakpoints did not have a positive impact on naïve yoked subjects. Only the rules generated after 570 trials of training had a positive impact on naïve subject performance. In a somewhat different manipulation, McGeorge and Burton (1989) asked subjects to write rules after 90 trials of training. When the researchers implemented these rules as computer programs, they found that a small number of them appeared able to outperform the subjects who had written them.

These findings suggest two important features of learning from outcome feedback in the SPF. First, even when explicitly directed to do so, the vast majority of subjects appear not to generate explicit hypotheses about the task, at least that would have a positive impact on naïve subjects' performance. This point accords well with Diehl and Sterman's (1995) observation that subjects quickly threw down their pads and pencils and trusted to intuition in learning to solve a somewhat more complicated dynamic decision task. Second, since the rules collected immediately after the breakpoints did not have a positive

[1] *Concordance* is specifically defined as the percentage of times that subjects gave the same response to the question as they had in training.

impact on naïve subjects' performance, the ability to generate such explicit rules is unlikely to precede improved performance.

The previous findings we have reported here lead to four important observations about learning in the SPF which are characteristic of more general results in the dynamic decision making literature. First, learning is approximate. Berry and Broadbent's (1987) findings suggest that, with training, subjects may become more accurate without eliminating all error in how they predict the SPF's behavior. In this vein, Dienes and Fahey's (1995) findings suggest that, with training, subjects also become more accurate but do not totally eliminate error in selecting actions to achieve their goal. Second, subjects' learning is local. Although Buchner *et al.* (1995) do not discuss near matches, their findings suggest that subjects' prediction performance is best for exactly the questions they have seen before. Dienes and Fahey (1995) find that, on workforce questions, nearness to past correct experience is critical to performance. Third, if learning is local and approximate, then subjects' ability to transfer knowledge they learn by doing the task will be best in areas very similar to their previous experience. Finally, Stanley *et al.'s* (1989) results suggest that improving performance during training in the SPF may not be the result of explicit hypothesis testing. In the next section, we provide a computational formulation of learning in dynamic decision tasks that can account for these observations.

3. A Computational Formulation of Learning

Our computational formulation combines a control theory framework with parallel distributed processing models to account for the local, approximate learning with graded transfer ability that subjects display in the Sugar Production Factory. In this section, we focus on each element of the computational formulation and how it contributes in general to the style of learning subjects display. In the next section, we generate and test predictions that the computational formulation makes about human subject learning in the task.

3.1. A Control Theory Framework

Brehmer (1990, 1992) proposes a control theory framework that characterizes decision makers in dynamic tasks as attempting to control a dynamic process in order to achieve desired outcomes. A problem for learning in this framework is that decision makers frequently do not receive direct feedback on what they should have done to achieve their desired outcome in a given situation. Therefore, Brehmer hypothesizes that decision makers' ability to adapt in dynamic decision-making tasks depends critically on their mental model of the environment for interpreting outcome feedback. In particular, Brehmer's (1990, 1995) laboratory subjects who possess less well-developed environment models have significant difficulty learning in more complex environments. However, Brehmer and other researchers have not specified the nature of decision makers' internal models, how decision makers learn these models while performing the task, and how decision makers use these models to improve their performance.

The issue of how internal models of the environment might be learned using outcome feedback was addressed by Jordan and Rumelhart (1992; Jordan, 1992, 1996) in motor learning, a standard application area for control theory. They used a strategy of dividing the problem of learning from outcome feedback into two interdependent subproblems: (1) learning the consequences of actions in given contexts and (2) learning what actions to take to achieve specific goals. In decision-making research, this division corresponds roughly to the difference between judgment and choice (Hogarth, 1981). With the learning problem subdivided this way, decision makers in dynamic environments may use knowledge about the consequences of actions to guide learning about what actions to take to achieve goals.

3.2. The Simulation Model

Building on this control theory hypothesis, we have constructed a computer simulation model (henceforth, the model) of how decision makers learn in dynamic decision tasks. In essence, the model learns by doing and deals with the two learning subproblems described above by placing two PDP submodels, *forward* and *action*, in series (see Fig. 1).

The submodels share a similar structure. The ovals in each submodel represent individual layers of simple processing units that take the outputs of other,

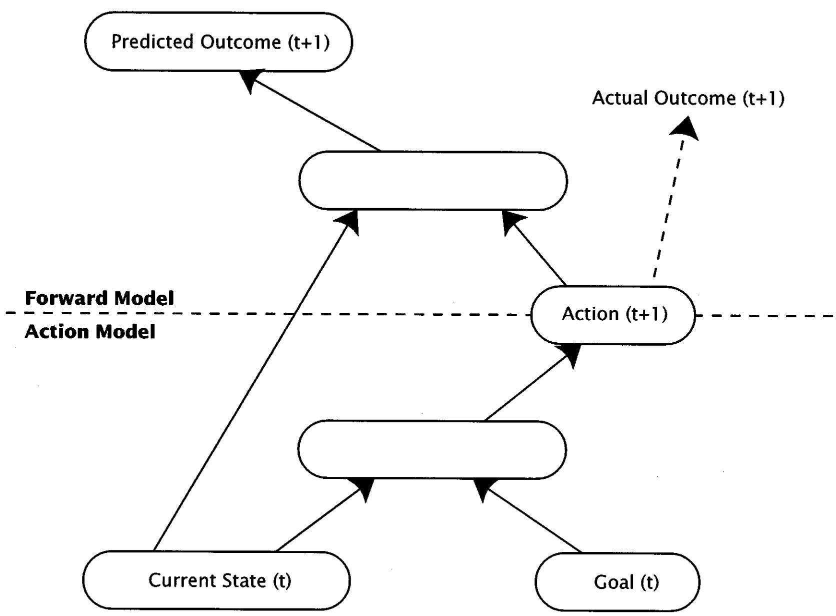


FIG. 1. A computational formulation of how decision makers learn from outcome feedback in dynamic tasks (derived from Jordan & Rumelhart, 1992). As described in the text, processing flows forward along the arrows from the input units through the middle layer to the output units. The action units are both an output of the action submodel and an input to the forward submodel. The environment processes the actions generated by the action units to produce the actual outcome.

connected processing units as their inputs, and aggregate them to produce their own output (Rumelhart, Hinton, & Williams, 1986b). The arrows represent weighted connections from all of the units in the layer at the tail of the arrow to all of the units in the layer at the head of the arrow. This connection structure constrains each submodel to map from inputs (earlier labeled ovals) through an internal representation of the inputs (unlabeled ovals) to outputs (later labeled ovals). Thus, by construction, all actions in the simulation model are based on immediate reactions to environmental situations. Based on our own (Gibson, Fichman, & Plaut, 1996) and others' observations (e.g., Klein *et al.*, 1993), this assumption appears to accord well with actual practice in field settings.

The way the individual units aggregate their input to produce outputs determines the type of mapping between input and output that each submodel produces. In the work reported here, the units in the input layers of both submodels simply take the values placed on them and pass them forward as their output to the units in the middle layer. The output layers consist of units that compute a weighted sum of their inputs from the middle layer to produce their outputs as shown in the following equation:

$$output_i = \sum_j w_{ij}\, output_j, \tag{3}$$

where $output_i$ is the ith unit's output, w_{ij} is the weight on the input connection from unit j's output to unit i's input, and $output_j$ is the output of unit j that i receives as input.

However, the units in the middle layers of each submodel pass the weighted sum of their input connections through a nonlinear logistic function to produce their outputs as shown in the following equation:

$$output_i = \frac{1}{1 + e^{-\Sigma_j w_{ij}\, output_j}}, \tag{4}$$

where $output_i$, w_{ij}, and $output_j$ are as in Eq. (3). This nonlinearity removes the constraint that the mapping computed by each submodel be linear.

For the simulations reported here, we use backpropagation to adjust the weighted connections between the units in the successive layers (Rumelhart *et al.*, 1986b). Backpropagation is a gradient descent learning algorithm that calculates how to adjust the connection weights to reduce the difference between environmental outcomes and the model's outputs at the output layer for each set of inputs. Our use of backpropagation coupled with the weighted summation process shown in Eq. (3) and (4) causes the two submodels to produce mappings that exploit the intercorrelational structure of their inputs and outputs (Rumelhart, Hinton, & Williams, 1986a).[2] Thus, each submodel is driven to produce

[2] Readers familiar with the backpropagation procedure will observe that the computation of the derivative for units using Eq. (3) and for units using Eq. (4) is different. Proper computation of these derivatives assures that backpropagation performs gradient descent (Rumelhart, Durbin, Golden, & Chauvin, 1995).

similar outputs for similar inputs in a way that reduces error at the output layer and that may be nonlinear.

Although the internal structure of the forward and action submodels is identical, their different inputs and outputs cause them to perform different tasks within the overall model. The action submodel takes as input the current state of the environment and the specific goal to achieve, and generates as output an action that achieves that goal. This action then leads to an outcome which can be compared with the goal to guide behavior. However, as we have noted, the environment does not provide direct feedback on how to adjust the action so as to improve the corresponding outcome's match to the goal.

Such feedback can be derived from the forward submodel. This network takes as input the current state of the environment and an action, and generates as output a predicted outcome. This predicted outcome is compared with the actual outcome to derive an error signal. Backpropagation is then used to adjust the connection weights between layers of the forward submodel to improve its ability to predict the effects of actions on the environment.

The forward submodel provides the action submodel with feedback for learning in the following way. The actual outcome produced by the action is compared with the goal to derive a second error signal. Backpropagation is again applied to the forward submodel (without changing its own connection weights) to determine how changing the action would change the error. This information corresponds to the error signal that the action submodel requires to determine how to adjust its connection weights so as to reduce the discrepancy between the goal and the actual outcome produced by its action. Note from this description that the accuracy of the forward submodel determines the quality of the error signal provided to the action submodel. Furthermore, the range of actions generated by the action submodel determines the range of the forward submodel's experience.

In summary, we have elaborated Brehmer's (1990, 1992, 1995) original proposal that an internal model of the environment plays a critical role in learning in dynamic decision tasks with three critical components. First, we hypothesize that learners acquire an internal model of the environment that may be characterized in terms of two PDP submodels, forward and action, that learn interdependently. Of these, the forward submodel, which learns how actions affect outcomes in given environmental contexts, appears to be the closest to Brehmer's original proposal. Second, we hypothesize that both the forward and action submodels are learned on-line as the task is performed using a method of error-correcting learning such as backpropagation. Third, the forward submodel aids the action submodel to learn by producing an error signal for the action submodel even though the environment does not directly provide this information.

4. COMPUTATIONAL PREDICTIONS AND EMPIRICAL EVIDENCE

In this section, we describe the computational simulation and subsequent human subjects experiment beginning with the general approach and then going to the particulars of each study.

4.1. General Approach

4.1.1. Task

As shown in Fig. 2, we presented human subjects with a graphical display of the SPF. This display was intended to supply two pieces of information about the current state of the factory which were likewise supplied to the model in the computational simulation. The first piece of information was the goal which was indicated by a horizontal bar spanning the production graph. The second piece of information was current production which was represented explicitly on the screen with a numerical value.

Both the model and subjects indicated the input for the next period in increments of fifty. In the case of subjects, they typed in the new input for the next period on the line provided near the bottom of the screen. This format closely resembles that used by several researchers (e.g., Berry & Broadbent, 1984; Buchner *et al.*, 1995; Dienes & Fahey, 1995).

4.1.2. Procedure

We provided learners (both the model and subjects) with three sessions of training on the SPF. At the end of the third session of training, we had them perform three additional activities. First, they answered control questions where they were given current production and had to supply a new control input to drive the factory to a production goal that they had not encountered during training. Second, learners answered prediction questions where they were given current production and a control input and had to predict the resulting production level for the next trial. Third, learners performed a transfer task where they had to achieve one of the goals that they had encountered in the control questions. The details of each stage are as follows.

Training. Similarly to Stanley *et al.* (1989), we trained our learners over

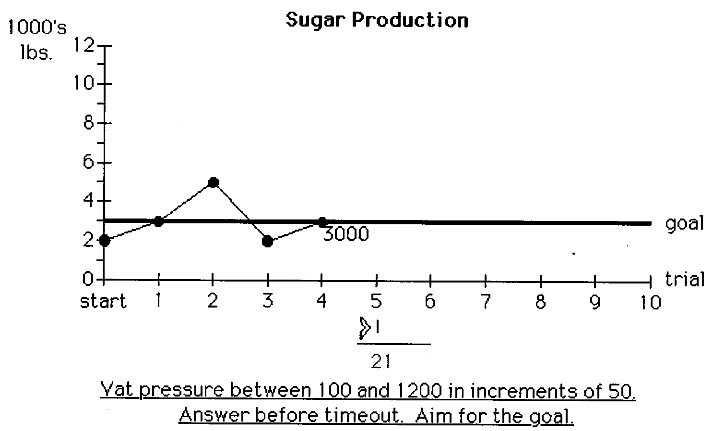


FIG. 2. Screen-shot of the Sugar Production Factory task as used in the human subjects study. At the beginning of each trial, subjects observed their performance relative to the production graph. Subjects indicated their control action (in this case, vat pressure) on the line below each trial. The sequence of numbers below this line represented a countdown clock that indicated to subjects how much time in seconds they had remaining in each trial.

600 trials. This training regimen is an order of magnitude longer than used in many studies (e.g., Berry & Broadbent, 1984, 1987; Buchner *et al.*, 1995). The long training period allowed us to examine learning beyond the initial stages exhibited in shorter experiments.

We broke the 600 training trials into three sessions of 20 sets of 10 trials. At the start of each set of 10 trials, production (P_t) was initialized to a random value and the goal was set to 3000 or 5000 lbs of sugar. Over the course of a session, learners experienced each goal a total of 10 times. We used multiple goals to provide learners with a broader basis for generalization from experience than that provided by using one goal during learning. Multiple goals have been used in this task by Buchner *et al.* (1995) and McGeorge and Burton (1989) but with a much shorter learning period.

Control questions. Immediately following the 200 training trials in session three, learners answered ten randomly ordered control questions. They were not given feedback about the correctness of their answers. The questions were evenly divided within learners between two production goals that were later used in the transfer task, 4000 lbs (*Near* to the training region) and 9000 lbs (*Far* from the training region) of sugar production. Learners had not seen these goals during training. Each question provided learners with the current production and the goal at the start of a set of ten trials in the same format that the learners had used to learn to control the factory. Learners were asked to provide a control input that would drive the factory to the indicated goal in the next trial. The *Near* goal of 4000 lbs was crossed with current production values 2000, 3000, 4000, 5000, and 6000 lbs of sugar. The *Far* goal of 9000 lbs was crossed with current production values of 7000, 8000, 9000, 10000, and 11000 lbs of sugar. In this way, the questions tested learners' ability to specify a control action that would keep them at or near goal performance once they had come within 2000 lbs. of a transfer goal that they had not seen before.

Prediction questions. Next, learners were presented with 18 randomly ordered questions that asked them to predict the effects of prespecified workforce values given different current production values at the start of a set of 10 trials. There were two sets of nine questions in the series. The first (*Near*) set consisted of current production values of 2000, 4000, and 6000 lbs of sugar crossed with control actions of 200, 400, and 600. The second set (*Far*) consisted of current production values of 7000, 9000, and 11000 lbs crossed with control actions of 700, 900, and 1100 to produce an additional nine scenarios. Both ranges of questions were designed to test learners' knowledge of the system's performance centered around the two transfer production goals of 4000 and 9000 lbs of sugar.

These questions were presented in a format as close as possible to the format that the learners used to learn the task. However, as described below, there were some differences in the format that human subjects used to answer these questions and the one they used for training.

Transfer. Finally, learners were asked to perform four sets of 10 trials of the task. They were assigned to either the *Near* condition with a production goal of 4000 lbs of sugar or the *Far* condition with a production goal of 9000 lbs of sugar. Starting production values for each set of 10 trials were chosen randomly without replacement from 3000, 6000, 7000, and 10000 lbs of sugar.

4.2. Computational Simulation

Here, we describe the computational simulation that generated predictions for the subsequent human subjects experiment. The predictions consisted of two components: (1) the pattern of performance in training, control, prediction, and transfer and (2) the level of performance in these activities.

4.2.1. Method

We instantiated the model (see Fig. 1) in the following way to learn to control the SPF as it performed the task. The goal production value was indicated as a real value on a single goal unit. The current production was represented as a real value on a separate input unit. Each of these inputs was scaled linearly to between 0 and 1. Finally, the middle layers in both the forward and action submodels each contained 30 units with their individual outputs computed by the logistic function in Eq. (4) scaled to ± 1. The number of middle layer units was established based on a series of pilot simulations intended to determine the minimum middle layer units required to learn a slightly more complex version of the task.

As described earlier, the forward and action submodels were trained with two different error signals. The predicted outcome generated by the forward submodel was subtracted from the actual (scaled) production value generated by Eq. (1) to produce the error signal for the forward submodel. The error signal for the action submodel was generated by subtracting the actual production generated by Eq. (1) from the goal level and backpropagating this error signal through the forward submodel as described earlier.

One training trial with the full model occurred as follows. The initial input values, including the goal, were placed on the input units. These then fed forward through the action middle layer. A single action unit using Eq. (3) took a weighted sum of the action middle unit activations, and this sum served as the model's indication of the workforce for the next time period. This workforce value was used in two ways. First, conforming to the bounds on workforce stipulated earlier, the value was used to determine the next period's production using Eq. (1). Second, the unmodified workforce value served as input into the forward submodel, along with the current production value. These inputs fed through the forward middle layer. A single predicted outcome unit computed a weighted sum of the forward middle unit activations using Eq. (1) and this sum served as the model's prediction of production for the next period. It is important to note that the forward and action submodels were trained simultaneously.

At the start of training, the connection weights of both the forward and action submodels were set to random initial values sampled uniformly between ±0.5. However, Berry and Broadbent (1984) observed that naive human subjects appear to adopt an initial "direct" strategy of moving workforce in the same direction that they want to move production. Our assumption was that this strategy resulted from prior experience with systems where it was adaptive. To approximate this prior experience, we pretrained our model for two sets of ten trials on a system in which change in production was directly proportional to change in workforce without lagged or random error terms. This pretraining biased our model toward ignoring the role of current production in producing the next period's production contrary to the relation stipulated in Eq. (1). Gibson and Plaut (1995) found that models with this bias better fit human data.

The use of initial random connection weights also caused each model to show individual learning characteristics, even after pretraining. Therefore, to get an accurate estimate of the abilities of the model, 40 instances (with different initial random weights prior to any pretraining) were trained. In all cases, the predictions for human performance attributed to the model are point estimates computed as averages of the performance of these 40 instances.

In the course of training, backpropagation (Rumelhart et al., 1986b) was applied and the weights of both the forward and action submodels were updated after each trial with a learning rate of 0.025 and no momentum.[3] We set the learning rate in such a way that the average training performance scored by the model closely approximated the average score of three human subjects in a pilot study.

We simulated the model's response to the tests after training as follows. For control questions, we placed values on the units representing the goal and current production for the model and measured the value on the action unit. For prediction questions, we placed the corresponding values on the model's units representing current production and action, we then interpreted the model's anticipated output unit as its prediction. Note that in this second variety of question, no value was placed on the goal unit. As described earlier, this unit does not affect processing in the forward submodel. Finally, the transfer task was simulated as another four sets of training with one of the two transfer goals. As described earlier for the general case, the four initial starting conditions were presented in random order without replacement.

4.2.2. Results

We tested each prediction using 1 *df* planned comparisons. To deal with the issue of nonindependence of repeated measures within models, we used Judd and McClelland's (1989, pp. 403–453) technique of orthogonal, polynomial contrast codes to create a single difference score for each model for each comparison. When dealing with proportions, we used arcsin transformations to account for non-constant variance.

[3] The learning rate parameter determines how quickly the weights change on each trial. Zero momentum means that there is no influence of past weight changes.

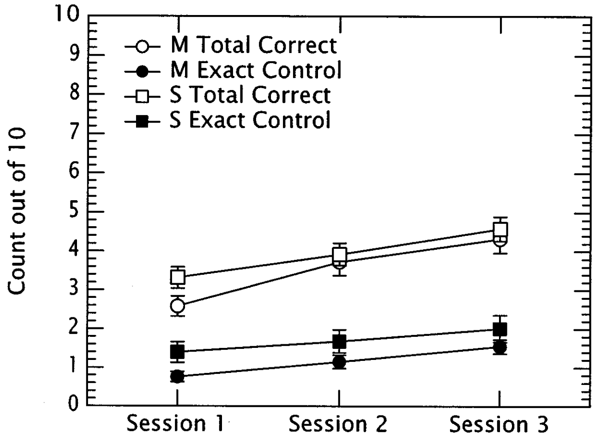


FIG. 3. Total correct and exact control for models and human subjects. Data items labeled S represent human subject data. Data items labeled M represent model data. The graph shows data from 24 human subjects and 40 models. The error bars represent one standard error for each data set. Human subject data are discussed in Section 4.3.

Model training. Model training results are presented in Fig. 3. The first set of predictions relates to Total Correct, the average number of trials within $\pm$ 1000 lbs of the production goal expressed as a count out of 10 (used in most SPF studies, e.g., Berry & Broadbent, 1984; Buchner *et al.*, 1995). The linear trend is highly significant ($t_{38} = 6.4270$, $p < .0001$). In spite of the non-linearity in the figure, the quadratic trend is not significant ($t_{38} = 1.8039$, $p = .0792$).

The second set of predictions in Figure 3 concerns Exact Control, the average number of trials that the learner specified the exact control action that would have been computed using Eq. (2), again expressed as a count out of 10. The linear trend for Exact Control is highly significant ($t_{38} = 5.8245$, $p < .0001$). There is no quadratic trend apparent in the figure ($t_{38} = 0.0264$, $p = .9791$).

The third predicted pattern of results for training is contained in Fig. 4 which shows Training Deviation tabulated by session. Training Deviation is

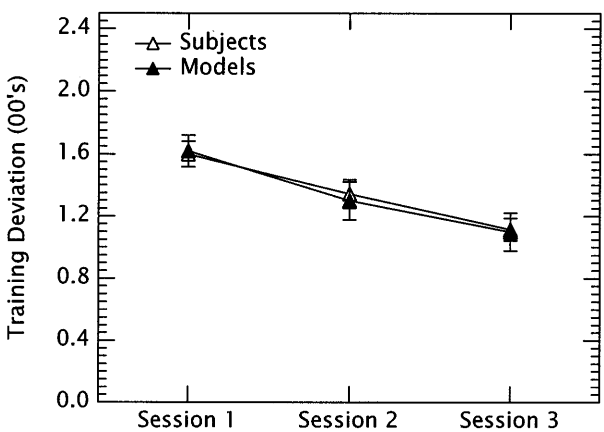


FIG. 4. Subjects' and models' reduction in training deviation. Data are from 24 human subjects and 40 models. The error bars represent one standard error for each data set. Human subject data are discussed in Section 4.3.

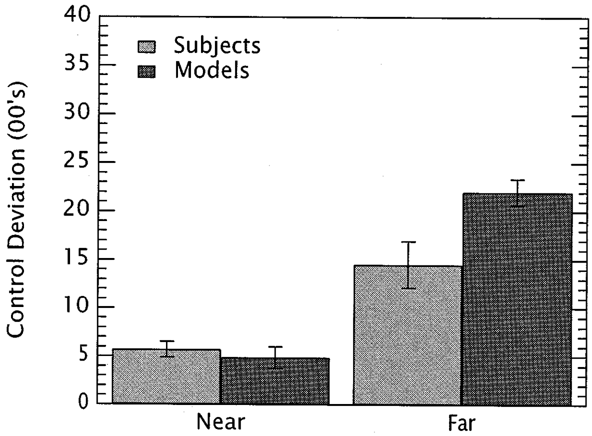
the average absolute deviation per set of ten trials of the learner's action from the exact action required to bring the factory to the production goal as computed using Eq. (2). This predicted pattern of results is different from the pattern related to Exact Control because it is possible to have a reduction in Training Deviation without a corresponding increase in Exact Control. As shown in the figure, the linear trend is highly significant (t_{38} = 7.3040, p < .0001). Again, in spite of the the nonlinearity apparent in Fig. 4, the quadratic trend is not significant (t_{38} = −1.7746, p = .0840). Subjects will decrease the discrepancy between their actions and those required to always bring the factory to within ± 1000 of goal production during training.

Control questions: action submodel knowledge at the end of training. In light of the training results, what is the extent of the knowledge the model gains about the system? Figure 5a shows average Control Deviation for the ten control questions. Control Deviation is the sum of the absolute difference between the learner's response and the response generated using Eq. (2). Note that the lowest average Control Deviation is in the *Near* column, where the transfer goal is nearest the training production goals of 3000 and 5000 lbs. The model predicts a significant positive difference for human subjects of 17.3 (t_{39} = 12.653, p < .0001) between the *Near* and *Far* conditions. This result illustrates the local nature of the model's learning. The model selects actions to achieve the production goals. As it becomes better at this task, it spends more and more time learning about system responses and refining its own responses near the production goals without regard to its general performance because it receives no error signal about that performance.

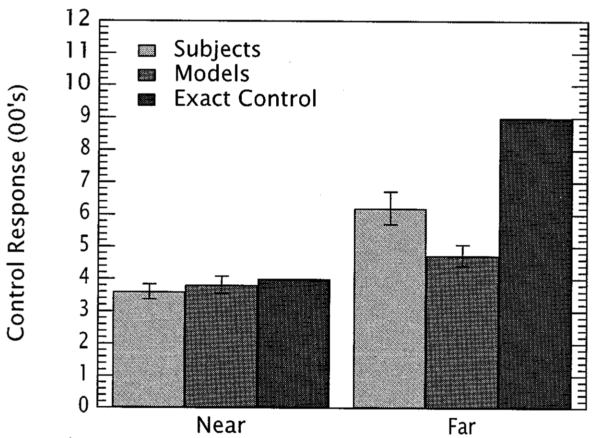
Nonetheless, the model does predict that subjects will adjust their responses for context as indicated by the significant difference in Control Response, the average of each learner's responses, between *Near* and *Far* questions (t_{39} = 5.8412, p < .0001). As seen in Fig. 5b, in the column labeled *Near,* the model's average response is very near the average of 400 produced using Eq. (2) and not significantly different from it (t_{39} = −0.7046, p = .4852). However, in the column labeled *Far,* the model's average response is much farther from the average of 900 produced using Eq. (2) and is significantly less than it (t_{39} = −13.0595, p < .0001). Human subjects will not adjust their responses sufficiently for the *Far* questions relative to what might be achieved using Eq. (2).

Prediction questions: forward submodel knowledge at the end of training. Figure 6a shows the model's average Prediction Deviation for *Near* and *Far* questions. Prediction Deviation is the sum of the absolute difference between the learner's response and the response generated using Eq. (2). The difference of 21.66 between the two columns is highly significant (t_{39} = 9.8913, p < .0001). As with the control questions, the model predicts that subjects will adjust their Prediction Response (average response over questions) for context as shown by the significant difference in this measure between *Near* and *Far* questions (t_{39} = 3.0466, p = .0041).

As shown in Fig. 6b, the difference in Prediction Deviation between *Near* and *Far* questions is that the model predicts that human subjects will not



(a) Average Control Deviation

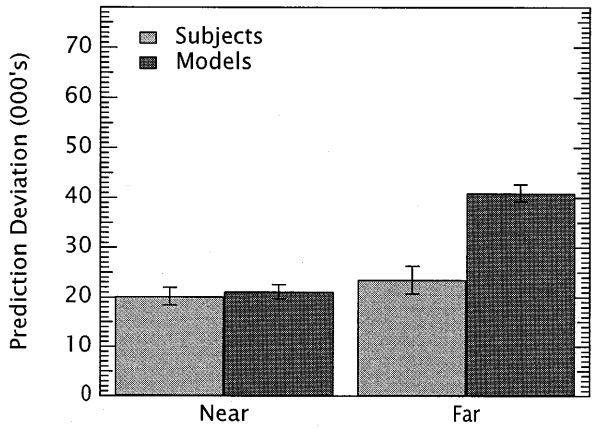


(b) Average Control Response

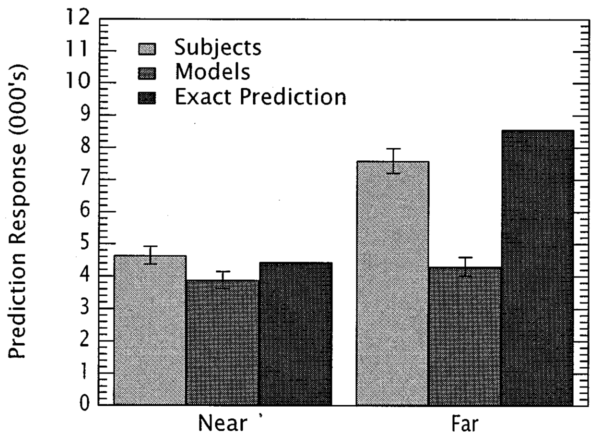
FIG. 5. Two charts representing control deviation and control response. In both charts, the columns labeled *Near* relate to a transfer goal of 4000 lbs of sugar. Those labeled *Far* relate to a transfer goal of 9000 lbs of sugar. Data are from 24 human subjects and 40 models. Error bars represent one standard deviation for each data set. Human subject data are discussed in Section 4.3.

adjust their responses sufficiently for the *Far* questions relative to what might be achieved using Eq. (1). In the column labeled *Near,* the model's average response is very near the average of 4,440 produced using Eq. (1) although significantly different from it ($t_{39} = -2.1546$, $p = .0374$). In the column labeled *Far,* the model's average response is much farther from the average of 8560 produced using Eq. (1) ($t_{39} = 14.8038$, $p < .0001$).

Transfer. Given the large differences in performance between *Near* and *Far* questions for both control and prediction questions, effective transfer performance with the *Far* goal over a small number of sets of trials would seem unlikely. As suggested by Fig. 7b, the difference for Total Correct between models in the *Near* and *Far* conditions is significant ($t_{38} = -3.5842$, $p = .0009$). *Near* subjects should outperform *Far* subjects. The linear trend ($t_{38} = 2.7177$,



(a) Average Prediction Deviation



(b) Average Prediction Response

FIG. 6. Two charts representing average prediction deviation and average prediction response. In both charts, the columns labeled *Near* relate to a transfer goal of 4000 lbs of sugar. Those labeled *Far* relate to a transfer goal of 9000 lbs of sugar. Data are from 24 human subjects 40 models. Error bars represent one standard error for each data set. Human subject data are discussed in Section 4.3.

p = .0098) is also significant as is the interaction between this trend and the transfer condition (t_{38} = 2.8391, p = .0072). *Far* subjects should improve over the four sets while *Near* subjects should not. As for the other trends, neither the quadratic trend (t_{38} = 1.4630, p = .1517) nor its interaction with the transfer condition (t_{38} = 1.6392, p = .1094) are significant. Additionally, neither the cubic trend (t_{38} = 0.4329, p = .6676) nor its interaction with transfer condition (t_{38} = −0.9502, p = .3480) are significant.

4.2.3. Summary of Results

We now review the computational simulation results in light of the approximate, local learning with graded transfer suggested by previous research using the SPF.

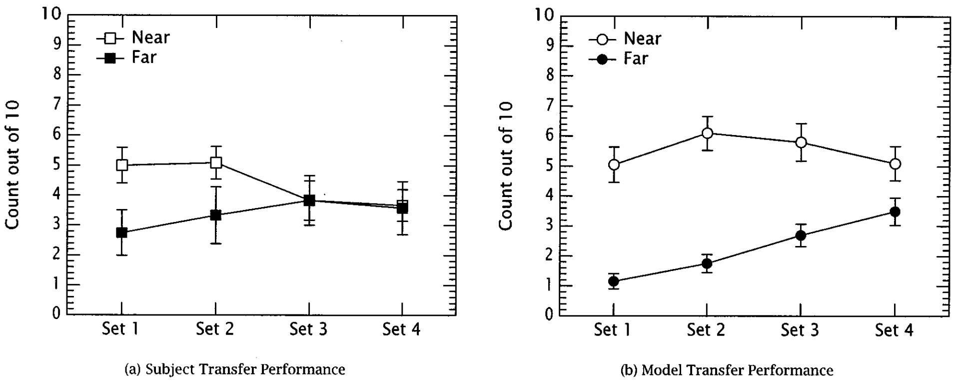


FIG. 7. Subject and model performance for the total correct measure in the *Near* and *Far* condition. Data are from 24 human subjects and 40 models. Human subject data are discussed in Section 4.3.

Approximate learning. As shown in Fig. 3, approximately two-thirds of trials counted in Total Correct during training result from inexact control actions. Although Exact Control increases as learning progresses and Training Deviation decreases, a large portion of performance results from actions that differ from those that would result from consistent application of Eq. (2).

In our formulation, learning is approximate for two reasons. First, the forward submodel only provides an approximate error signal to the action submodel. Second, the form of error-correction we are using reduces error more slowly as learning progresses, leaving learning in some sense always approximate. In the SPF, both of these effects are compounded by the random element (ϵ) in Eq. (1) that distorts feedback.

Local learning. As shown in Figs. 3 and 4, models appear to have the most accurate knowledge of the underlying system in the areas near where the production goals are during training. By accurate knowledge, we mean the least amount of error in choosing actions or predicting system response. In general, the difference in performance between *Near* and *Far* questions appears to derive from the learning bias inherent in the model learner's architecture. The model is a goal-driven learner. As it improves, it learns to stay in a region near its training goals. It is not adjusting its connection weights relative to some global measure of performance but only relative to its current goal based on information presented in the last trial.

Graded transfer ability. In the first transfer set, performance is graded by the proximity of the transfer goals to the original training goals. This result derives from the approximate, local learning exhibited by the model during training. The model only acquires approximate knowledge of the system. This knowledge is most accurate in the area of its greatest experience. The transfer goal of 9000 lbs. was chosen to be the maximum distance possible from the training region with the constraint that it be more than 2000 lbs. from the maximum production of 12,000.[4] Therefore, it is not surprising that performance is lower for this goal.

In spite of this approximate and biased learning, the model exhibits improving performance for the goal of 9000 lbs. to the point that performance for this goal is very close to performance for the goal of 4000 lbs. by the fourth set. This quick adaptation derives from the model's ability to continue learning from feedback in new contexts.

4.3. Human Subjects Experiment

In this section, we describe a computer-controlled experiment designed to test the simulation model's predictions.

4.3.1. Method

The human subjects experiment was implemented in cT (Sherwood, 1994) on an Apple Macintosh computer using a 256 color display. Except as noted below for the prediction questions, subjects indicated their control inputs and question answers using an Apple extended keyboard.

We now consider features of the human subjects study that were different from the simulation study.

Time pressure. Initial pilot studies indicated to us that, while most subjects seemed to perform in the SPF taking little time to mentally explore or consider options, a few seemed to be mentally exploring options at great length. The best evidence of this difference in approach is the distribution of times it took subjects to complete one session of the task. This distribution ranged from 20 min to 1 h. We used time pressure to remove the option of spending long amounts of time mentally exploring or reconsidering options. Time pressure used in this way is a general feature of dynamic decision tasks (Brehmer, 1990, 1992, 1995; Klein *et al.*, 1993).

We added time pressure to the SPF by limiting subjects to three seconds per trial to take a control action.[5] We enforced this time limit by providing subjects with a counter directly beneath the input area for each trial that told them how much time they had left. If subjects did not specify a control action within this time, a bright red screen covered the production graph and a beep sounded repeatedly until the subject responded. These two events were designed to stop subjects from considering the task and to encourage them to answer as quickly as possible.

To further limit the chances to mentally review the task, after each set of 10 trials, the screen used to present the task to subjects was blanked and

[4] This choice ensured that subjects could not achieve an average score of 3.33 Total Correct by simply maximizing their input at each trial.

[5] We derived this time limit based on analogy to real decision makers' performance in the field of credit collections, a dynamic decision task. Credit collectors make decisions on average every 2.0 s during a conversation (Gibson *et al.*, 1996). We increased the limit to 3.0 s based on experience during pilot studies.

subjects were allowed up to 10 s to initiate a new set of 10 trials at their discretion. Once 10 s had passed, a new set of 10 trials automatically began. Blanking the screen inhibited subjects from actively contemplating their actions from the previous set of 10 trials. Limiting time between sets of 10 trials further constrained subjects' ability to contemplate the task without the aid of the display.

The cover story. By construction, our simulation model is not sensitive to the cover story used in the task. However, the addition of time pressure forced us to construct a cover story so that the task would make sense to human subjects. To this end, subjects were told that they were process control engineers at a sugar production factory. Their job was to set the pressure on a sugar-refining vat to achieve a target production level. A safety mechanism on the vat required that the engineer specify the pressure every 3 s, otherwise pressure might rise to dangerous levels. After 3 s, a red safety screen would descend, and an alarm would sound.

Procedure. Subjects for the experiment consisted of 24 Carnegie Mellon students who were paid $18 for their participation. These subjects were randomly assigned to either the *Near* (4000 lbs) or *Far* (9000 lbs) transfer condition upon beginning the experiment.

The experiment was conducted in three sessions over three consecutive days. The procedure differs from the computational study because subjects had a 24-h break between sessions, whereas models did not. This break is potentially significant because it could have led to performance differences between models and subjects. For instance, the break could have led to forgetting on the part of subjects. On the other hand, some pilot subjects reported actively replaying the events of the experiment in their mind between sessions, possibly leading to improved performance. We did not confine the experiment to one session to remove these possible influences on subjects' performance because we found subjects becoming fatigued between 200 and 400 training trials, and we wanted to maintain the extent of training.

At the end of training in the third session, the control questions were presented using the same display subjects had used to learn the task. Each set of questions was presented in random order, and subjects did not receive feedback about the correctness of their response. Again the three second time limit was enforced.

After the control questions, subjects were asked the prediction questions. Subjects had not been asked prediction questions during the experiment and therefore were not used to giving this type of response. However, given the structure of our simulation model, we believed subjects could anticipate how the system would respond to their actions and observe the difference between their anticipation and the actual outcome. Based on this assumption, we asked subjects to use their computer mouse to indicate the area where they thought production would fall given current production and the workforce that we had specified in the question. A similar procedure was used by Wolpert *et al.* (1995) when they measured people's ability to predict the visual position of their hands after movement using only proprioceptive feedback. Subjects seemed able to perform our procedure readily. Their mouse clicks were captured by the computer and used to calculate the estimated production level to the nearest thousand. This value was then displayed to the subject. Finally, the transfer task immediately followed the questions.

4.3.2. Results

Timing. Our model does not account for how subjects will react to different levels of time pressure nor the timing of their responses during the experiment. The goal of adding time pressure elements to the task was to force subjects to react quickly to the situation at hand. Subjects averaged 2.07 ($SE = 0.078$), 1.76 ($SE = 0.07$), and 1.69 ($SE = 0.067$) per decision for sessions one, two, and three, respectively, during training. Additionally, time per decision showed a significant decreasing linear trend ($t_{22} = -4.7647$, $p < .0001$) and quadratic trend ($t_{22} = -3.1578$, $p = .0046$) across subjects. Not only were subjects able to initially make the three second time limit, but they became faster as training progressed.

Subjects also made the three second time limit in answering the control questions (mean = 2.71, $SE = 0.102$), the prediction questions (mean = 2.42, $SE = 0.093$), and the transfer task (mean = 1.81, $SE = 0.095$). The rest of this section focuses on the account the model is able to give of the human data.

Training. Figure 3 allows comparison of human subject and model learning over the three sessions of training. Looking at Total Correct, human subjects show a significant, positive linear trend in performance ($t_{22} = 5.6885$, $p < .0001$) without a significant quadratic trend ($t_{22} = -0.2440$, $p = .8095$). Both results conform to the model's predictions. Furthermore, the significant improvement across sessions replicates Stanley *et al.'s* (1989) results for this length training regimen.

Turning to Exact Control, human subjects again show a significant, positive linear trend in performance ($t_{22} = 4.5025$, $p = .0002$), without a significant quadratic trend ($t_{22} = -0.3440$, $p = 0.7341$). Again, both results conform to the model's predictions.

Looking again at Fig. 3, the model and human subject performance means appear very close for both the Total Correct and Exact Control measures. We performed post-hoc comparisons for both variables between models and subjects. For this type of comparison, the Bonferroni adjustment is inappropriate. By increasing the size of the confidence interval around each subject mean, it raises the likelihood of false confirmations that the model's predictions lie in this confidence interval. Furthermore, one may want to use a narrower band than the 95% confidence interval to make this comparison (Stasser, 1988). Therefore, for this type of post-hoc comparison, we simply report the unadjusted t and p values.

For Total Correct, there is a significant difference between subjects and the model ($t_{62} = 2.0566$, $p = .0439$) in the first session. For sessions two and three there is not a significant difference between subject and model means ($t_{62} =$

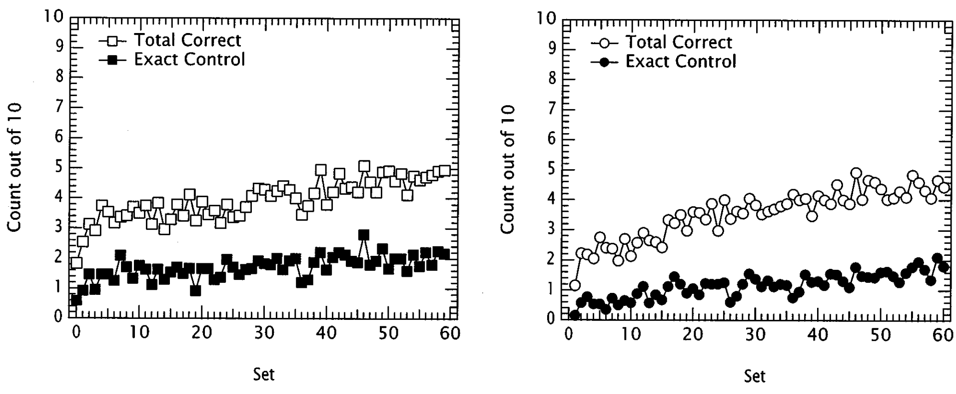
0.6309, p = .5304 for session two; t_{62} = 0.7618, p = .4491 for session three). For Exact Control, there is a significant difference between subjects and models for session one (t_{62} = 2.9793, p = .0041) and session two (t_{62} = 2.0285, p = .0486). There is not a significant difference for session three (t_{23} = 1.6006, p = .1146).

We also tested whether each subject mean differed significantly from the performance of ten that could be achieved with consistent application of Eq. (2). The three session means all differed significantly from this level of performance ($t_{23} = -35.3624$, $p < .0001$ for session one; $t_{23} = -30.1679$, $p < .0001$ for session two; $t_{23} = -24.2775$, $p < .0001$ for session three).

Figure 8 presents average human subject and model learner performances for the 60 training sets for the variables Total Correct and Exact Control. We tested the correlation between average human and model performance over the 60 sets using a Bonferroni adjustment on the confidence level. For Total Correct, the correlation of 0.80 is highly significant (t_{58} = 10.2142, $p < .0001$). For Exact Control, the correlation of 0.62 is also highly significant (t_{58} = 5.9499, $p < .0001$). Thus, model performance tracks human performance well across training sets for both Total Correct and Exact Control.

Figure 4 shows human subjects' average Training Deviation per set of ten trials for each session of training. As predicted by the model, human subjects show a significant linear trend in performance ($t_{22} = -8.1159$, $p < .0001$) with no quadratic trend ($t_{22} = -0.2764$, p = .7848). Finally, as suggested by the figure, none of the human subject means is close to being significantly different from the model's predictions in post-hoc comparisons ($t_{62} = -0.1422$, p = .8874 for session one; t_{62} = 0.2526, p = .8014 for session two; and t_{62} = 0.0972, p = .9229 for session three).

However, there is a highly significant difference between average subject performance and the performance of zero deviation that could be obtained through consistent application of Eq. (2) (t_{23} = 19.019, $p < .0001$ for session one; t_{23} = 14.6204, $p < .0001$ for session two; t_{23} = 15.4941, $p < .0001$ for session three). As predicted by the model, subjects reduce the difference between their



(a) Subject Performance (b) Model Performance

FIG. 8. Improvement in performance with training broken down for the two measures total correct and exact control. The graph represents data from 24 human subjects and 40 models.

actions and the action produced using Eq. (2) without bringing the difference to zero.

Figure 9 compares human subjects' average Training Deviation by set during training with that predicted by the model. The correlation of 0.86 is highly significant after Bonferroni adjustment ($t_{58} = 12.8532$, $p < .0001$) again indicating that the model tracks human performance well across sets.

Finally, as noted earlier, many studies (e.g., Berry & Broadbent, 1984; McGeorge & Burton, 1989; Buchner *et al.*, 1995) show that human subjects improve their performance over a relatively short 60 to 90 trials. We tested whether models and subjects improved performance over 60 trials in our training regimen by dividing the trials into three sets of twenty and testing the appropriate contrasts. For Total Correct, both models and subjects showed significant linear trends in performance ($t_{38} = 3.0399$, $p = .0043$ for models; and $t_{22} = 3.2046$, $p = 0.0041$ for subjects) and insignificant quadratic trends ($t_{38} = 0.099$, $p = 0.9217$ for models; and $t_{22} = 0.2937$, $p = .7717$ for subjects). The model shows significant performance improvements over 60 trials, and human subjects display a similar trend in performance improvement over the same number of trials.

Control questions. Figure 5a shows subjects' Control Deviation measured in hundreds for *Near* and *Far* control questions. As predicted by the model, the difference in Control Deviation is significant for subjects ($t_{23} = 5.2182$, $p < .0001$). We also performed post-hoc tests to determine whether the Control Deviation predicted by the model for each set of questions fell within the confidence interval of the mean performance across subjects. For *Near* questions, the average deviation across subjects is not significantly different from the model's prediction ($t_{62} = 0.5178$, $p = 0.6064$), but this deviation is significantly different from zero ($t_{23} = 7.0732$, $p < .0001$), the result that would have been obtained with consistent use of Equation 2. For *Far* questions, the average deviation across subjects is significantly different from the model's prediction ($t_{62} = -2.9456$, $p = .0045$) and again from zero ($t_{23} = 6.061$, $p < .0001$).

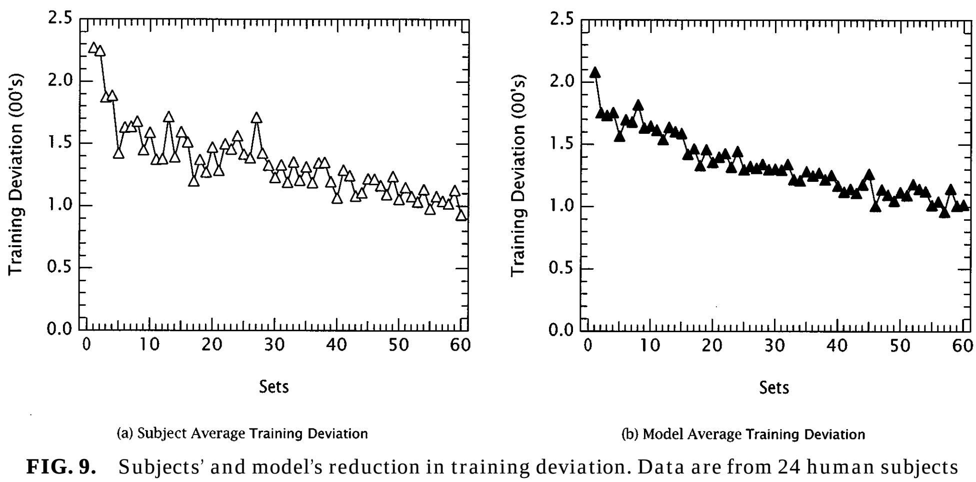


FIG. 9. Subjects' and model's reduction in training deviation. Data are from 24 human subjects and 40 models.

Recall from Fig. 5 that the model predicts a positive difference between the average Control Response provided in the *Far* and *Near* conditions. This difference is significant across human subjects (t_{23} = 7.0635, $p < .0001$). We also compared the average Control Response taken in the *Near* and *Far* questions across subjects with the model's predictions and with the average control response generated by using Equation 2 consistently. For the *Near* questions, subjects' answers were not significantly different from the model's predictions ($t_{62} = -0.5337$, $p = .5955$) nor from the average response obtained by using Equation 2 ($t_{23} = -1.6636$, $p = .1098$). For the *Far* questions, subjects' average Control Response was significantly different from both the model's predictions (t_{62} = 2.5579, p = .013) and from the average Control Response given by consistently using Eq. (2) ($t_{23} = -5.5509$, $p < .0001$). Thus, while the average responses produced by models, subjects, and consistent application of Equation 2 are not significantly different for *Near* questions, they are so for *Far* questions. For *Far* questions, subjects average Control Response lies between that produced by the model and that produced by consistent application of Equation 2 but is much closer to the average produced by the model.

The training goals of 3000 and 5000 lbs of sugar are much nearer together than the goals in the *Near* and *Far* control questions. Do models and subjects respond differently to the smaller differences in goals during training? Both models (t_{39} = 3.4647, p = .0013) and subjects (t_{23} = 8.0963, $p < .0001$) show a significant positive difference between responses to the 3000 and 5000 lbs training goals. This result augments our findings by showing that subjects and models are sensitive to even the smaller differences in goals present during training. Nevertheless, the key result of this analysis remains that neither subjects nor models adjust their response sufficiently to goals and production values that are far from the training region.

Finally, many previous studies of performance in the SPF (e.g., Berry & Broadbent, 1984) have shown little or no correlation between performance during training and correctness in answering post-task questionnaires. An exception to this result occurs when the questions are presented in the same modality used during training and correctness is measured as whether the response is in the proper direction (e.g., Berry & Broadbent, 1987). Similarly to Berry and Broadbent (1987), we asked the control questions in the same modality subjects had used to perform the task and used a deviations measure that was more sensitive than simply determining that subjects were or were not exactly correct. Both models ($t_{38} = -6.9924$, $p < .0001$) and subjects ($t_{22} = -2.2101$, $p = .0378$) show negative correlations of -0.75 and -0.43 respectively between average Total Correct during training and total Control Deviation on the control questions. This is the direction that would be expected if better training performance were correlated with better question answering. This result accords well with Berry and Broadbent's (1987) finding on question answering. However, although in the right direction to suggest a cause for this result, there is no significant correlation between training performance and a smaller difference in Control Deviation between *Near* and *Far* questions (correlation = −0.39, t_{38} = −0.3937, p = .6960 for models; correlation = −0.32, t_{22} = −1.5610, p = .1328 for subjects).

Prediction questions. Figure 6 shows subjects' Prediction Deviation measured in thousands for *Near* and *Far* prediction questions. Contrary to the model's prediction, the difference in Prediction Deviation, while in the correct direction, is not significant for subjects (t_{23} = 1.6784, p = .1068). We also performed post-hoc tests to determine whether the Prediction Deviation predicted by the model for each set of questions fell within the confidence interval of the mean performance across subjects. For *Near* questions, the average Prediction Deviation across subjects is not significantly different from the model's prediction (t_{62} = −0.4204, p = .6757). The average Prediction Deviation is significantly different from 0 (t_{23} = 11.3823, $p < .0001$), the result that would have been obtained by consistent use of Eq. (1) before application of the random element. For *Far* questions, the average Prediction Deviation across subjects is significantly different from the model's prediction (t_{62} = −5.6021, $p < .0001$) and again from 0 (t_{23} = 19.9428, $p < .0001$). However, as is apparent in Fig. 6a, the average Prediction Deviation across subjects for *Far* questions is not significantly different from what the model predicts for the *Near* questions (t_{62} = 0.8391, p = 0.4046).

Recall from Fig. 6 that the model predicts a small positive difference between the average Prediction Response given for the *Far* and *Near* questions. This difference is significant across human subjects (t_{23} = 5.022, $p < .0001$). We also compared the average Prediction Response given in the *Near* and *Far* questions across subjects with the model's predictions and with the average response generated by using Eq. (1). For the *Near* questions, subjects' average Prediction Response was not significantly different from either that of the model (t_{62} = 1.908, p = .061) or from the average answer obtained by using Eq. (1) (t_{23} = 0.7253, p = 0.4756). For the *Far* questions, subjects' average Prediction Response was significantly different from the model's Prediction Response (t_{62} = 6.9312, $p < .0001$) as well as from the average response given by using Eq. (1) (t_{23} = −2.5547, p = .0177).

Again, we tested the correlation between training performance and question answering with the expectation that better training performance would be correlated with better question answering. As for the control questions, both models (t_{38} = −5.6683; $p < .0001$) and subjects (t_{22} = −2.301; p = .0312) show strong negative correlations of −0.67 and −0.50, respectively, between average Total Correct and total Prediction Deviation, confirming our expectation. Again, better training performance is not significantly correlated with a smaller difference between *Near* and *Far* questions for Prediction Deviation, although it is in the right direction for models (correlation = −0.17, t_{38} = −1.0852, p = .2847) but not for subjects (correlation = 0.01, t_{22} = 0.0703, p = .9446).

Thus, the model did predict the non-zero Prediction Deviation for both *Near* and *Far* questions as well as the difference in average Prediction Response between *Near* and *Far* questions. Furthermore, the correlation between training and question answering is significant for both subjects and models. However, the difference for subjects in Prediction Deviation between *Near* and *Far* questions, though in the right direction, was not significant. Furthermore, the model's predictions for average Prediction Response were significantly different from subjects' averages for both *Near* and *Far* questions.

Transfer. Figure 7 contrasts average human and model performance in the transfer task using the Total Correct measure for the *Near* (4000 lbs) and *Far* (9000 lbs) transfer conditions. Recall that the model makes essentially two predictions about human performance. First, there will be a significant negative difference between subjects in the *Far* and *Near* transfer conditions. Although the difference is in the right direction for human subjects in the two conditions, it is not significant ($t_{22} = -1.2134$, $p = .2379$). Second, the model predicts an interaction between transfer condition and the evolution of performance using the Total Correct measure over sets. Learners in the *Far* condition should show an improvement in performance over the four sets and learners in the *Near* condition should not. This prediction is in the right direction for the interaction between transfer condition and the linear trend across the four sets but is not significant ($t_{22} = 1.8691$, $p = .0750$). Additionally, the significant linear trend predicted by the model is not found in human subjects ($t_{22} = -0.5751$, $p = 0.5734$). As with the model, human subjects do not show significant quadratic ($t_{22} = -0.7144$, $p = .4825$) or cubic ($t_{22} = 0.2709$, $p = .7890$) trends nor do they show significant interaction effects between these trends and transfer condition ($t_{22} = -0.5215$, $p = .6072$ for quadratic; $t_{22} = -1.0274$, $p = .3154$ for cubic).

The interaction between transfer condition and the linear trend in performance across sets observed in the model suggests that, using the Total Correct measure, differences in performance between the subjects in the *Near* and *Far* conditions should be more easily observed in the earlier sets of the transfer task (see Fig. 7a). The difference between transfer conditions is not significant for either set one or set two after adjustment ($t_{22} = -2.3380$, $p = .0435$ for set one; $t_{22} = -1.6017$, $p = .128$ for set two).

We also performed post-hoc comparisons of average human subject performance with the model's predictions. All of the model's predictions lay within a 95% confidence interval of the relevant subject means.

Note, however, in comparing Fig. 7a and 7b, that the model does not account for an important trend in the human subject data. Performance of subjects in the *Near* transfer condition decreases after set two, and performance for subjects in the *Far* condition fails to improve after set three. We believe that this trend can be explained by the fact that subjects knew that the transfer task was the last task they would perform in the experiment and so left off trying at the end of the experiment. Such "horizon" effects have been noted in other dynamic decision tasks (Sterman, 1989).

As with the control and prediction questions, we tested the correlation between training and transfer performance for both models and subjects. Again, both models (t_{38} = 6.1624; $p < 0.0001$) and subjects (t_{23} = 8.2091; $p < .0001$) showed positive correlations of 0.73 and 0.87 respectively between better training and better transfer performance. Furthermore, this relationship holds when we control for whether learners are in *Near* or *Far* conditions using an analysis of covariance ($\beta = 0.8151$, $t_{36} = 8.0922$, $p < .0001$ for models; $\beta = 0.7998$, $t_{20} = 5.8288$, $p < .0001$ for subjects).

For the transfer task, the model provides, as its strongest account of human performance, the interaction between transfer condition and the linear trend in performance. Subjects in the *Far* condition start significantly lower than subjects in the *Near* condition but improve to parity with these subjects by the third set. Furthermore, both the model and subjects show significant correlations between better training and transfer performance.

Summary and Discussion of Human Subjects Results

The model provides a good account of human subject performance for training, control questions, and transfer. For training, the model correctly predicted the strong linear trends in performance for Total Correct, Exact Control, and Training Deviation. Post-hoc comparisons showed the model's average performance by session to be insignificantly different from that of human subjects for two out of three sessions using Total Correct, for one out of three sessions using Exact Control, and for three out of three sessions using Training Deviation. Furthermore, over the 60 training sets, model and human subject performance were highly and significantly correlated for all three measures.

All three of the measures improved during training, but none of them reached the level that could be achieved using Eq. (2). Like models, subjects only learned to achieve approximate control of the system. As we noted earlier, this result is consistent with prior results using the SPF.

Subject performance in the control questions provides further evidence for human subjects' approximate learning while also demonstrating the local nature of their learning. First, as predicted by the model, the difference between subjects' average Control Deviation and 0 is highly significant for both *Near* and *Far* control questions. Second, as predicted by the model, the difference in Control Deviation between *Near* and *Far* questions is also highly significant and positive. After 600 trials of training, subjects have not acquired, or at least are not consistently using, an accurate understanding of the mechanism underlying the SPF's performance in answering these questions. On average, subjects perform better in situations that most closely match their previous training experience and worse in situations that are further from it, demonstrating the local nature of subjects' knowledge. These results are again consistent with those found in previous studies (Buchner *et al.*, 1995; Dienes & Fahey, 1995; Marescaux *et al.*, 1989).

In spite of the approximate, local nature of subjects' knowledge demonstrated by these findings, subjects, like models, are sensitive to context. The difference between Control Response between *Near* and *Far* questions is significant for human subjects. However, like models, subjects tend to underadjust for the *Far* control questions.

Finally, as regards the control questions, both models and subjects show correlation between better training performance and question answering, consistent with Berry and Broadbent's (1987) finding. For models, the result can only occur because, during training, some models have better acquired the structure of the task than others. This observation also holds for the correlations exhibited by models between better training performance and better performance in prediction and transfer. For these same correlations exhibited by human subjects, the explanation that applies to models is plausible. Additionally, there may be explanations relating to motivational and other possible factors. We did not test these other explanations for human subjects and so cannot say definitively whether the model's explanation holds for them.

While the results for transfer performance do not reach the same level of statistical significance as those described for training and control questions, they do provide additional support for local, approximate learning with the addition of graded transfer. Like the model, in both the *Near* and *Far* transfer conditions, subjects' mean Total Correct is significantly different from the performance that would have resulted from consistent application of Eq. (2) demonstrating that subjects have acquired only approximate knowledge about the SPF. Second, while the contrast between conditions using Total Correct was not significant for human subjects, it was in the direction predicted by the model. Furthermore in this regard, a post-hoc comparison between the means for the first set of the transfer task showed a difference in the direction predicted by the model, although this difference was not significant after adjustment of the confidence interval. Both of these last two pieces of evidence point to the greater validity of human subjects' knowledge in the locality of their training experience. Third, the model correctly predicted the direction of the interaction between transfer condition and the linear trend in the Total Correct measure. This illustrates the graded nature of transfer. Subjects are able to adapt to the *Far* goal, but they start at a lower level of performance and require more time to adapt than subjects with the *Near* goal.

However, the model does not provide a good account of two aspects of human performance. First, while the model does predict the difference between *Near* and *Far* conditions for control questions, it underpredicts subjects' level of performance in the *Far* condition of both tasks. The model does not correctly predict the level of Control Deviation in the *Far* condition which is larger than for human subjects, nor does it correctly predict the level of Control Response which is less for models than human subjects. In short, the model appears less adaptive to the *Far* questions than are human subjects.

The second area where the model does not predict human performance well is for the prediction questions. The model predicts a difference in Prediction Deviation between *Near* and *Far* questions that is not statistically significant but in the predicted direction for subjects and a difference in Prediction Response between *Near* and *Far* questions that is significant. It also predicts a non-zero deviation for both *Near* and *Far* questions for subjects. However, the model fails to capture the level of Prediction Response for either the *Near* or *Far* conditions and the level of Prediction Deviation indicated for models in the *Near* condition is indistinguishable from that of subjects in *Far* condition.

The two shortcomings we have just described serve as an indication of areas for future research. Overall, the model's approximate, local learning with graded transfer provided a strong, predictive account of subject performance in training, control questions, and transfer where both subjects and models consistently used the same modalities across tasks. Our results concerning approximate, local learning in these tasks are consistent with prior empirical results using the SPF. Our transfer results are an extension of these prior results.

5. GENERAL DISCUSSION

In this article, we have presented evidence that a computational model that instantiates approximate, local learning with graded transfer provides a good account of how subjects learn on-line from outcome feedback in the SPF, a simple dynamic task. We base this conclusion on the model's ability to predict subjects' performance during training and on two subsequent tests of their ability to generalize, the control questions and the transfer task. We now explore the limitations of our efforts and discuss two alternative approaches to understanding human performance before concluding on our own approach's merits.

5.1. Limitations of Our Approach

There are three principal limitations of our approach. First, our model does not provide a good account of how subjects perform after training in prediction questions. One possibility is that the prediction task differed from the training task sufficiently that subjects simply abandoned the knowledge they had gained during training and began guessing based on one of the cues. In other words, our manipulation failed. In this regard, it is worth noting that Wolpert *et al.* (1995) demonstrate that a biased forward model provides a good account of human subjects' ability to predict the hand location that results from attempts to move their arms given proprioceptive feedback alone. The difference between their approach and ours is that their subjects provided the input for which they predicted the consequence. This difference in manipulation may have ensured that their subjects used the same knowledge that they normally would when moving their arms. We are actively looking into ways to adapt this approach to our SPF experiments.

The next two limitations to our approach touch directly on its applicability to real-world dynamic tasks. The first of these is that our approach only covers on-line learning from outcome feedback. We undertook specific steps during training to restrict subjects to this mode of learning and justified these steps based on the observation that there frequently is not much spare time in dynamic task environments. However, it is clear, even in dynamic task environments, that decision makers do have respite for reflection, if at no time else, during their off-hours. During this time, decision makers might review their previous actions and the resultant outcomes with an aim toward improving their actions. We could easily implement this extension in our model by replaying some of the training data. However, we deliberately limited this possibility in both the computational and human subjects experiments because it is difficult to control the extent or the relevance of subjects' thoughts without the impetus of time pressure driving them to stay focused on the task.

The third limitation to our approach is that all of the relevant data for making a decision must be immediately available in the environment. There are frequently delays between actions and outcomes such that the action producing a given outcome is no longer immediately available to be related to that outcome once it is known (Brehmer, 1995). While decision makers display some capacity for adaptation in such environments, this capacity is limited (Brehmer, 1995; Sterman, 1994). We are currently investigating ways to extend our approach to provide an account of human performance in such delayed feedback environments.

5.2. Alternative Approaches

Two alternative approaches to ours have been explored for explaining performance in the SPF. First, Stanley *et al.* (1989) have examined the possibility that subjects are using explicit hypothesis testing to generate their performance. Recall that in Stanley et al.'s *original learners* condition, they concluded that breakpoints, after which subjects' performance began to improve at a significantly faster rate had not been preceded by a communicable insight and were therefore not likely the result of explicit hypothesis testing.

More generally, Stanley et al.'s (1989) findings relate to the twin issues of explicit and implicit learning, the focus of a large number of studies using the SPF (e.g., Berry & Broadbent, 1984, 1987; McGeorge & Burton, 1989). When learners performance improves, and they are *explicitly aware* of task structure and how they are using it to make decisions, they are said to be engaged in explicit learning (Cleeremans, 1993). Evidence of explicit learning includes subjects' ability to verbally communicate how they are making decisions and their ability to provide accurate, relevant information about the task (Berry & Broadbent, 1984; McGeorge & Burton, 1989; Stanley *et al.*, 1989). When learners' performance improves and they are not explicitly aware of task structure, as measured by their ability to provide useful verbal information about the task, they are said to be engaged in implicit learning.

Our model is composed of two mental submodels, neither one of which depends on explicit awareness of task structure as it learns (see Cleeremans & McClelland, 1991 and Cleeremans, 1993 for a fuller discussion of PDP models in this regard). However, even with this assumption embedded in our model, we are not aware of results conclusively ruling out that subjects use some form of explicit learning (perhaps poorly communicable). This said, based on the pattern of results we obtained in training, the control questions, and the transfer task, it would seem that any such explicit learning mechanism would have to account for the approximate, local learning with graded transfer that we found in average subject responses. This statement is nontrivial because forms of explicit learning such as insight learning lead to the induction of precise, globally accurate rules concerning the system under study or rules that are obviously wrong (e.g., Kaplan & Simon, 1990).

The second alternative approach has been adapted from Logan (1988) by Dienes and Fahey (1995) to explain learning in the SPF over 80 trials with one goal. In Dienes and Fahey's approach there are two cognitive mechanisms within the learner competing to provide control inputs at each trial. The first, termed *the algorithm* by Logan (1988), is a fixed body of general knowledge that comes into play when the learner encounters a situation that has not been seen before. In Dienes and Fahey's implementation, *the algorithm* provides a model of the assumptions naive subjects bring to the SPF. The second cognitive mechanism in Dienes and Fahey's approach is responsible for learning. It is a lookup table that adds context-action exemplars every time an action leads to within 1000 of the production goal.

Dienes and Fahey's approach has provided a compelling account of learning and performance in their training regimen. The chief difference between our approach and that of Dienes and Fahey is apparent in how we dealt with the initial biases subjects bring to the SPF. Our assumption was that these biases were due to subjects' experience with other superficially similar systems that behaved differently from the SPF. Therefore, instead of having recourse to an additional cognitive mechanism such as *the algorithm*, we pretrained our models in a system where these biases were adaptive, as a proxy for subjects' prior experience. Thus, we were able to provide an account of subject behavior in our training regimen that relied on only one performance and learning mechanism. Of course, our results do not rule out that multiple performance or learning mechanisms may be needed to explain behavior in other tasks or even in this task under other circumstances.

Although Dienes and Fahey's (1995) approach does not address the issue of prediction, we believe it could have replicated at least our training results. In their approach, we note that approximate performance results from the rule used to match context to action. Local learning results from their use of a lookup table to encode specific context-action pairs. The issue that remains to be explored is how Dienes and Fahey's approach would have performed in the control questions and the transfer task. Due to important differences between the information we provided to learners and that provided by Dienes and Fahey, informed speculation on this issue is difficult.

Dienes (personal communication) has suggested that one way to address these two issues would be to query naive subjects about the task implementation we used. Dienes and Fahey's model could then be run in our training and testing regimen allowing us to make a direct comparison between the different predictions for the control questions and the transfer task.[6] Predictions concerning response times and attentiveness to different elements of context, two features that distinguish Dienes and Fahey's approach, could also be tested. We believe that a direct comparison between our approach and that of Dienes and Fahey is a useful path for future work.

5.3. Conclusion

Our results suggest that the behavior that decision makers display in dynamic decision tasks can be captured by a formulation where two PDP submodels learn on-line, joined in series. One submodel learns to take actions to achieve goals. The other learns the consequences of actions in different contexts and guides the first model's learning. The fact that both submodels are PDP models provides a mechanism for learning from feedback *and* graded generalization from the specific examples seen in training (Rumelhart *et al.,* 1986a). In this way, our formulation provides both a structure and a single learning mechanism for how decision makers might form and use approximate, locally valid knowledge of task structure while performing dynamic decision tasks.

Our simulation model improves its performance within a given range of objectives during training. However, after training, transfer of this knowledge to achieving a new objective in the same task was graded by the nearness of the new objective to those used during training. Both of these patterns of results were confirmed in our human subjects experiment, and they accord well with the observation that subjects do not adjust their behavior on-line in response to feedback in a way that fully takes into account task structure (Brehmer, 1995; Sterman, 1994; Kleinmuntz, 1993). Given these results, our formulation presents an important step toward a process theory of how decision makers learn on-line from outcome feedback in dynamic tasks.

REFERENCES

Berry, D. C. (1991). The role of action in implicit learning. *Quarterly Journal of Experimental Psychology,* **43A**(4), 881–906.

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.

Berry, D. C., & Broadbent, D. E. (1987). The combination of explicit and implicit learning processes in task control. *Psychological Research,* **49,** 7–15.

Berry, D. C., & Broadbent, D. E. (1988). Interactive tasks and the implicit-explicit distinction. *British Journal of Psychology,* **79,** 251–272.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues.* Hillsdale, NJ: Erlbaum.

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

[6] Alternatively, Dienes and Fahey's procedure or a mix of the two procedures could be used. An advantage to using 600 training trials, as we did, is that it enables comparison beyond 80 trials.

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, & J. Funke (Eds.), *Complex problem solving: The European perspective.* Hillsdale, NJ: Erlbaum.

Buchner, A., Funke, J., & Berry, D. C. (1995). Negative correlations between control performance and verbalizeable knowledge: Indicators for implicit learning in process control tasks? *Quarterly Journal of Experimental Psychology,* **48A**(1), 166–187.

Cleeremans, A. (1993). *Mechanisms of implicit learning.* Cambridge, MA: The MIT Press.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General,* **120,** 235–253.

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. (1990). *Implicit concept formation.* Ph.D. thesis, University of Oxford.

Dienes, Z., & Fahey, R. (1994). *The role of implicit memory in controlling a dynamic system.* Unpublished manuscript.

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.

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

Funke, J. (1992). Dealing with dynamic systems: Research strategy, diagnostic approach, and experimental results. *The German Journal of Psychology,* **16**(1), 24–43.

Gibson, F. P. (1996). The sugar production factory—a dynamic decision task. *Journal of Computational and Mathematical Organization Theory,* **2**(1).

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., & 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: Erlbaum.

Hayes, N. A., & Broadbent, D. E. (1988). Two modes of learning for interactive tasks. *Cognition,* **28**(28), 249–276.

Hogarth, R. M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. *Psychological Bulletin,* **90,** 197–217.

Jordan, M. I. (1992). Constrained supervised learning. *Journal of Mathematical Psychology,* **36,** 396–425.

Jordan, M. I. (1996). Computational aspects of motor control and motor learning. In H. Heuer, & S. Keele (Eds.), *Handbook of perception and action: Motor skills.* New York: Academic Press.

Jordan, M. I., Flash, T., & Arnon, Y. (1994). A model of the learning of arm trajectories from spatial deviations. *Journal of Cognitive Neuroscience,* **6**(4), 359–376.

Jordan, M. I., & Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science,* **16**(3), 307–354.

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

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology,* **22,** 374–419.

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.

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

Marescaux, P. J., Luc, F., & Karnas, G. (1989). Modes d'apprentissage selective et non-selective et connaissances acquises au controle d'un processus: Évaluation d'un modèle simulé. *Cahiers de Psychologie Cognitive Européenne,* **9,** 239–264.

McGeorge, P., & Burton, A. M. (1989). The effect of concurrent verbalization on performance in a dynamic systems task. *British Journal of Psychology,* **80,** 455–465.

Rapoport, A. (1975). Research paradigms for studying dynamic decision behavior. In D. Wendt, & C. Vlek (Eds.), *Utility, probability, and human decision making.* Dordrecht-Holland: D. Reidel Publishing Company.

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: Erlbaum.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986a). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & the PDP research group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations* (pp. 318–362). Cambridge, MA: MIT Press.

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

Sanderson, P. M. (1990). *Implicit and explicit control of a dynamic task: Empirical and conceptual issues* (Tech. Rep. No. EPRL-90-02). Urbana, IL 61801: Department of Mechanical and Industrial Engineering, University of Illinois at Urbana-Champaign.

Sherwood, J. N. (1994). *cT version 2.0: Reference manual.* Raleigh, NC: Physics Academic Software.

Squire, L. R., & Frambach, M. (1990). Cognitive learning in amnesia. *Psychobiology,* **18**(1), 109–117.

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.

Stasser, G. (1988). Computer simulation as a research tool: The discuss model of group decision making. *Journal of Experimental Social Psychology,* **24**(5), 393–422.

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.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). Forward dynamic models in human motor control: Psychophysical evidence. In *Advances in Neural Information Processing Systems 7.* Cambridge, MA: MIT Press.

Received: July 8, 1996