



## Instance-based learning in dynamic decision making

Cleotilde Gonzalez\*, Javier F. Lerch, Christian Lebiere

*Department of Social and Decision Sciences, Carnegie Mellon University,  
5000 Forbes Avenue, Pittsburgh, PA 15213, USA*

Received 16 January 2001; received in revised form 28 October 2002; accepted 10 January 2003

## Abstract

This paper presents a learning theory pertinent to dynamic decision making (DDM) called instance-based learning theory (IBLT). IBLT proposes five learning mechanisms in the context of a decision-making process: instance-based knowledge, recognition-based retrieval, adaptive strategies, necessity-based choice, and feedback updates. IBLT suggests in DDM people learn with the accumulation and refinement of instances, containing the decision-making situation, action, and utility of decisions. As decision makers interact with a dynamic task, they recognize a situation according to its similarity to past instances, adapt their judgment strategies from heuristic-based to instance-based, and refine the accumulated knowledge according to feedback on the result of their actions. The IBLT's learning mechanisms have been implemented in an ACT-R cognitive model. Through a series of experiments, this paper shows how the IBLT's learning mechanisms closely approximate the relative trend magnitude and performance of human data. Although the cognitive model is bounded within the context of a dynamic task, the IBLT is a general theory of decision making applicable to other dynamic environments. © 2003 Cognitive Science Society, Inc. All rights reserved.

**Keywords:** Dynamic decision making; Instance-based learning; Cognitive modeling; Decision making; Water purification plant

## 1. Introduction

Dynamic decision making (DDM) has been characterized by multiple, interdependent, and real-time decisions, occurring in an environment that changes independently and as a function of a sequence of actions (Brehmer, 1990; Edwards, 1962). Kerstholt and Raaijmakers (1997)

\* Corresponding author. Tel.: +1-412-268-6242; fax: +1-412-268-6938.  
E-mail address: [conzalez@andrew.cmu.edu](mailto:conzalez@andrew.cmu.edu) (C. Gonzalez).

review laboratory studies in DDM and summarize the main findings. Most previous research is devoted to understanding why people perform poorly in dynamic tasks rather than understanding the *process* of decision making. To learn the causal and temporal relationships of decisions and outcomes, dynamic tasks should allow decision makers lengthy practice. However, most research does not allow extended practice. We have performed several behavioral studies to understand how decision-making skills are developed overtime in dynamic situations (Gonzalez, submitted for publication). In these studies, we have measured the decision maker's actions against heuristics overtime. We have found that decision makers improve their performance by following heuristics less closely and more inconsistently. Experienced decision makers show a lower fit to heuristics and higher standard deviation compared to their own behavior at the beginning of their practice. Our interpretation is that overtime decision makers increasingly use their accumulated knowledge to make decisions and take advantage of their prior knowledge. Based on these results, we have proposed that the most likely learning mechanism in DDM is the acquisition and retrieval of decision instances or examples. This proposition is supported by theories of decision making under uncertainty (Gibbs & Schmeidler, 1995, 2000) as well as by observations of decision makers acting on time-constrained real world situations (Klein, Orasanu, Calderwood, & Zsombok, 1993; Pew & Mavor, 1998; Zsombok & Klein, 1997). Decision makers must become proficient in DDM tasks by taking advantage of domain-specific knowledge through practice. As decision makers observe situations and make decisions, they become more proficient at determining which decisions work best under specific situations.

In this paper, we propose a set of learning mechanisms applicable to dynamic decision environments. The instance-based learning theory (IBLT) proposes that in DDM situations people learn by accumulation, recognition, and refinement of instances. Instances contain information on the decision-making situation, the action, and the result of a decision. Then, we present the IBLT implementation into the cognitive model, CogIBLT, in a dynamic decision task. Through a series of experiments the mechanisms proposed by IBLT are compared to human data, presenting an analysis of model and human data at a micro level, and exploring how the model and presumably humans learn in dynamic environments. Finally, we present a discussion of the results and our conclusions.

## 2. Learning and skill acquisition in dynamic decision making

Psychology is full of learning theories. These theories have been developed under different views, such as, implicit and explicit learning (e.g., Berry & Broadbent, 1984; Merrill, Sun, & Peterson, 2001), learning from examples and by doing (e.g., Anderson et al., 1981; Simon & Amazez, 1979; Simon & Zhu, 1988), and deductive and inductive learning (e.g., Medin, Wattenmaker, & Michalski, 1987). Simon and Langley (1981) defined learning as "a process that modifies a system as to improve, more or less irreversibly, its subsequent performance of the same task or of tasks drawn from the same population." Since there are many ways in which the human cognitive system may be modified, they suggested that it is more realistic to have theories of learning *mechanisms* rather than theories of learning. Simon and Langley provided a taxonomy of learning mechanisms in complex tasks including: the "*Knowledge base*," characterized by the accumulation of knowledge in declarative form; "*Recognition*,"

presented as the ability to discriminate among familiar classes of objects; "*Strategies*," referring to adaptive production systems; and "*Evaluation Functions*," related to the assessment of different alternatives to control the continuation of search in problem solving. In this section, we focus our review on instance-based learning theories, and present evidence of the relevance of these theories to real world and complex tasks. The research introduced next contributes to the *knowledge base* mechanism as well as to other learning mechanisms towards an integrated theory (IBLT) for decision making in dynamic environments.

The chunking theory originally proposed by Chase and Simon (1973) and derived from EPAM<sup>1</sup> suggests that learning occurs by the accumulation of chunks in long-term memory and that experts in a particular domain recognize chunks and hold links to them in short-term memory. This theory applied to chess playing, suggests that skill develops mainly through recognition of features and familiarity of chunks. Experts search very selectively, using environmental cues to guide their attention and achieving great computational efficiency, while novices engage in more exhaustive search (Chase & Simon, 1973; Simon & Gobet, 1996). Heuristics or rules of thumb allow a skilled player to be restricted to a small tree of possibilities. Those heuristics are determined by the recognition of familiar patterns of chunks, giving access to information stored in memory about possible good actions. More recently, this theory has been revised to suggest that highly skilled people use other long-term memory structures (templates) in addition to chunks in short-term memory (Simon & Gobet, 1996). Experts then, may not only retrieve larger chunks, but templates filled with chunks in the slots. Simon and Gobet (1996) use the concept of templates to explain the ability of chess masters to recall the main features of many of the games they have played. Chess playing is a complex task, and it can be dynamic if it is played under time constraints. Seen from the perspective of one player exogenous and independent environment are formed by the actions of the opponent. The chunking theory contributes to the *recognition* and *strategies* mechanisms within IBLT.

Hintzman (1984, 1986) proposes an instance-based recognition model where he defines a memory trace as a list of task features within the record of an experience. This theory explores the acquisition of abstract concepts from examples. Knowledge accumulates in secondary memory (SM), and traces are retrieved from SM by the communication of a probe.<sup>1</sup> This theory was implemented into a computational system called MINERVA 2. Learning in MINERVA 2 consists of copying the features of an experience into the trace structure using a probabilistic method. Despite Hintzman's focus on simple tasks (paired-associate learning of lists) the *recognition* mechanism proposed in memory trace theory may be relevant to dynamic decision tasks. In contrast to the chunking theory, in Hintzman's theory there is no need for a separate generic memory structure, but retrieval from SM is based on the similarity to the probe and SM traces at the time of retrieval.

Medin and Schaffer (1978) propose a theory of classification and context, based on stored examples, to deal with ill-structured natural environments. This model states that people learn to classify objects based on the retrieval of stored examples. The theory assumes that classification of a stimulus is determined on the basis of its similarity to stored category examples. Nosofsky (1984) extends the context model to modeling choice and similarity. In particular, he integrates the concept of similarity into the literature of stimulus similarity and proposes a mathematical function to represent the computation of similarity. Similarity is a concept on which many instance-based learning theories rely. Medin, Goldstone, and Gentner (1993)

review the concept of similarity and its use both theoretically and empirically. They propose that similarity varies with experience and with context, and similarity judgments depend on the details of the comparison process. This is a construct used in many other models of categorization and instance-based learning (Cohen & Nosofsky, 2000; Fortus, Gentner, & Law, 1994; Palmer, 1997). We believe that similarity plays a key role in *recognition, strategies, and evaluation functions* learning mechanisms in real world and complex tasks.

The Instance Theory of Automization, proposed by Logan (1988), provides a model of skill acquisition based on retrieval of examples from memory. A key characteristic of this theory is a transition from algorithm-based to instance-based performances. Learners move from general algorithms to specific solutions as they gain experience in a task. These solutions are retained in memory and retrieved when the same problems are confronted. Logan proposes that retrieval of previous instances is based on a race between memory retrieval and algorithm execution: whichever finishes first, controls the response. He assumes that people switch to memory retrieval over rules because of the accumulation of instances and the higher probability of instances winning the race. Another important characteristic of this theory is that the process of encoding examples is a consequence of attention. The acquisition of expertise is due to the increasing reliance on memory retrieval rather than computational methods. Multiple laboratory studies support the notion that automatic performance relies on memory retrieval rather than on rule application (Lassaline & Logan, 1993; Logan, 1990, 1992; Logan & Klapp, 1991; Schneider & Shiffrin, 1977). Logan's theory, however, does not rely on the concept of similarity as part of memory retrieval, but rather suggests that only instances *identical* to the present items can be retrieved from memory. Nosofsky and Palmer (1997) extend Logan's model to propose that memory retrieval is similarity based. Responses are determined by the similarity of a case to a previous set of responses (Palmer, 1997). Nosofsky and Palmer propose the Exemplar-Based Random Walk (EBRW). This mechanism assumes a similarity-based race of exemplars to be retrieved. Logan's theory and its extensions contribute to the *strategies and the evaluation function* learning mechanisms.

Fields other than psychology also investigate the use of experience in the form of instances or examples to understand and plan for novel situations. In the decision sciences, "Case-Based Decision Theory" (CBDT) presents a very elegant mathematical theory of decision making under uncertainty (Giboa & Schneider, 1995, 2000). CBDT proposes that decision-making situations are stored in the form of examples where people choose successful acts in cases similar to those they recall from past experience. However, this theory has neither been verified nor validated with human data. In artificial intelligence (AI) a large number of studies, mainly based upon computational modeling, investigate the process of case acquisition and retrieval in complex tasks. Case-Based Reasoning (CBR) proposes that skills in complex tasks are developed by the accumulation of cases containing the information that characterizes the state of the world, the solution to that problem, and the outcome (describing the state of the world after the case occurred) (Reisbeck & Schank, 1989). CBR demonstrates that a way to solve novel problems is to adapt previously successful solutions to similar problems (Ram, 1993). AI research also studies hybrid approaches to learning. For example, Domingos (1996) attempts to combine rule induction and instance-based approaches generalizing rules from instances in a gradual process of classification. Also, reinforcement learning addresses how computational agents discover actions that yield most reward by interacting with the environment.

Computational research of reinforcement learning is now an important research topic in AI (Sutton & Barto, 1998). All learning research in AI, from CBR to reinforcement learning, provides great insight into how human learning might occur in complex environments. However, AI is concerned with the computational accuracy and functional utility of the algorithms produced. Most AI studies are not concerned with how well they represent and reproduce human learning.

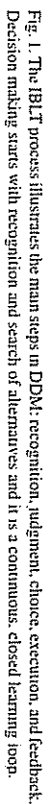
Instance-based learning theories bring together the following characteristics: accumulation of examples in memory through training and task repetition, development of pattern recognition and selective alternative search, similarity-based memory retrieval, gradual withdrawal of attention while increasing memory retrieval, and transition from rule-based to exemplar-based performance. These theories, however, have been tested in simple and static tasks, some focusing more on automatic rather than deliberate actions. Despite the lack of real world validation of these instance-based theories, we believe they are all relevant to real world complex tasks like DDM. In many real world situations, researchers have observed and suggested instance-based behavior. For example, in military and war decision making, it has been suggested that instance-based methods are important components of learning (Pew & Mavor, 1998). In military conditions, cues for making decisions are uncertain and conflicting, the results of taking different alternatives are difficult to imagine and little is known about the current state of the world. Frequently, military commanders substantiate their decisions on experience, comparing previous situations to the present case (Pew & Mavor, 1998). In other real world tasks, it has been suggested that complex cognitive skills are acquired by the transition from analytical to intuitive thinking (Dreyfus, H. L., 1997; Dreyfus, H. L. & Dreyfus, S. E., 1986). In Naturalistic Decision Making (NDM), researchers examine what people do in real world situations such as healthcare, aviation, and fire fighting (Klein et al., 1993; Zsombok & Klein, 1997). It has been observed that experts use recognition and satisficing instead of analytical and optimizing strategies to make decisions.

In summary, very little is known about how decision skills develop in dynamic and complex tasks. We believe that instance-based learning theories have characteristics that apply well to DDM. Research suggests that in DDM people accumulate knowledge in the form of examples, retrieve those examples by selecting among familiar categories, and evolve from computational-based strategies to instance-based retrieval. Presented next is the initial development of an IBLT compiling several learning mechanisms into a framework for decision making in dynamic environments.

### 3. Instance-based learning in dynamic decision making

This section presents a theory containing the learning mechanisms suggested by Simon and Langley (1981) in the context of dynamic decision processes. It is called IBLT because the main knowledge element of IBLT is an instance and because decision making in dynamic environments involves learning.

Similar to a case in CBR, in IBLT an *instance* is defined as a triplet with situation, decision, and utility (SDU) slots. A *situation* is described as a set of environmental cues; a *decision* represents the set of actions applicable to a situation; and *utility* is the evaluation of the goodness of a decision in that particular situation. To illustrate the concept of SDU, imagine yourself



IBLT is not only about accumulation of instances (SDUs), but it is a compilation of learning mechanisms as proposed by Simon and Langley (1981). The IBLT mechanisms are:

- **Feedback updates.** Method to update the utility of SDUs and maintain the causal attribution of results to actions.

aggregated utility value from past experiences. Next, a decision point comes into place: to search for more alternatives or to execute the current best alternative. The answer to this choice affects the search for more alternatives or to execute the current best alternative. The answer to this choice affects the search for more alternatives or to execute the current best alternative.

of the SDUs provides a better representation of the "goodness" of an action. Below are the explanations of each of these steps in detail.

### 3.1. Recognition

gradual increase of reliance on memory and decrease on attention (Carr, McCauley, Sperber, & Parmalee, 1982; Marcel, 1983).

the field of decision sciences, Gilboa and Schneider (1997) also extended their original CBDT to take into account evaluation of acts depending on past effectiveness of *similar* acts.

similar situations closely match the cues of the current situation. Important cues from the environment stand out in the recognition process because they resemble cues in previous

instances. This guides attention and produces selective behavior. Recognition is based on attention, guided by previous knowledge and determined by instance similarity. Similarity may include both objective and subjective factors. It is calculated from the comparison of a current situation to the past experience based on either objective factors, subjective factors or both.

A situation may be considered atypical when no similar situations can be retrieved from memory, for example, when novices start to learn a task. In atypical situations decision makers do not know what cues of the task are important, therefore, predefined goals, non-contextual knowledge, and heuristics may guide attention focus. Decision makers confronted with atypical situations take longer to complete the recognition process because they have to pay attention to different parts of the environment with no context-based guidance. Recognition of typical situations is faster because previous knowledge is used to perceive the important cues and to ignore irrelevant factors. Finally, with practice, and after seeing similar situations, decision makers retrieve previous knowledge of situations similar to the current one and abstract the common features from the past to direct the attention to the important cues in the present.

### 3.2. Judgment

After recognizing a situation as typical or atypical, decision makers evaluate the accuracy of a possible action in a particular situation (e.g., determine utility). IBLT proposes two procedures for judgment: for atypical situations decision makers use heuristics while for typical situations they use previous knowledge (accumulated SDU instances). IBLT also proposes that decision makers adapt their strategies from heuristic-based judgments to instance-based judgments.

In atypical situations, decision makers use heuristics to evaluate a decision's potential success. Judgments may be based on the given instructions, the goal, environmental cues, or even random heuristics. In complex tasks such as DDM, people use heuristics that are very information selective (Gigerenzer & Todd, 1999; Payne, Bettman, & Johnson, 1993). Also, by practicing the task, decision makers take advantage of their experience and learn from the outcomes of previous decisions (Logan, 1992; Nosofsky & Palmeri, 1997). Most directly in DDM, decision makers decide when to intervene according to the changing situation itself. The time heuristic recommends making a decision according to the time left. This heuristic is basically the earliest time rule in operations management and can be applicable to any dynamic situation. In DDM, decision makers should learn to adjust the moment of intervention to the time left (Kerstholt & Raaijmakers, 1997).

In typical situations, IBLT proposes that decision makers determine the utility of an action by combining the utility from similar instances generated in the past. Retrieval of memory knowledge according to the probe similarity is well supported by instance-based learning theories. Likewise in CBR, a common retrieval algorithm, the "nearest neighbor" approach retrieves the past instance that is most similar to the current situation (Watson, 1995). In the field of decision sciences, Gilboa and Schneider (1995, 1997) argue that the evaluation of an act is a weighted sum of the similarity between problems, current problems and those stored in memory, and acts, both current and past. They indicate that retrieval utility evaluation is derived from all relevant cases to the decision at hand rather than from the nearest neighbor alone. In IBLT, we rely on the concept of "activation" associated with each SDU to determine which of the similar instances will be combined to calculate the utility of the current situation.

Activation, defined by Anderson and Lebiere (1998), is the reflection of the instance usefulness in the past and the relevance of that instance to the current context. IBLT proposes that all active instances similar to the current situation are retrieved from memory. Next, the utility values from all those past similar situations are combined into a new evaluation of the current situation. New SDU instances are produced and the utility is the accumulation of previous knowledge.

### 3.3. Choice

After the judgment process, decision makers determine the best course of action. Rational theories of choice assume that all alternatives relevant to the choice are known, precise, consistent, and stable while the best of all alternatives is determined by the decision maker (March, 1994). This is clearly not how DDM works. IBLT suggests an intermediate strategy between the *optimizing* and the *satisficing* strategies of choice. IBLT proposes that alternatives are evaluated one by one and after each evaluation a choice is made between searching for more alternatives and executing the current best alternative. Similar to the satisficing strategy proposed by Simon, IBLT suggests that the evaluation of more alternatives is determined by the decision maker's necessity level. In Simon's satisficing strategy, the alternative executed is the first option that works (based on an aspiration level), regardless of time or events in the environment. In IBLT, the necessity level indicates the need to make a decision, measured by subjective or objective factors such as preferences, or time left. In DDM, a major determining factor of "satisfaction" is the time left to make a decision and the interruptions due to exogenous events. As suggested from experimental studies of DDM, people fail to determine and adjust the moment of intervention in a dynamic task because accurate timing requires extensive knowledge about the causal and temporal relations between situations and outcomes (Kerstholt & Raaijmakers, 1997). IBLT suggests the perception of time left (or sense of urgency) evolves overtime. In a dynamic environment, unexperienced decision makers might see every situation as urgent, reacting too early. With practice in the task, decision makers should learn the temporal relationship of events and outputs and react more closely to the moment that would produce best performance.

The necessity level determines the number of alternatives evaluated before a decision is executed. If there is no time left, the decision maker will execute the current best alternative. IBLT predicts that, if there is enough time, decision makers will follow an exhaustive evaluation of alternatives, selecting in fact the best one of all (optimizing strategy). During the evaluation process decision makers keep in mind the SDU with the best utility value, so that they may interrupt the evaluation process at any moment and execute the current best option.

The recognition-judgment cycles continue until an alternative is selected. The execution process simply indicates the actual implementation of the decision involved in the SDU selected. Executing a decision modifies the environment and modifies the SDU's stored in memory to indicate the alternative selected.

### 3.4. Feedback

To improve performance, the decision maker must be able to detect the results from the decisions made and feedback provides this knowledge. However, in DDM feedback is usually

delayed and it is hard to make a connection with the actions that produced such feedback. Interestingly, instance-based learning models provide little if any information on how feedback might be accounted for in learning.

IBLT suggests that decision makers use feedback to refine the SDU instances. When an SDU is created, the utility value is a prediction of the result of an action in the current conditions, but the output from the action is unknown until feedback is received. With knowledge of results, decision makers may update the utility value on the original SDUs. We believe decision makers re-evaluate how accurate decisions are, according to feedback from the environment. This upgrading process makes good SDU instances more likely to be retrieved in the future. On the other hand, poor outcomes should result in the decision maker downgrading the utility in the SDU instances making poor instances less likely to be retrieved in the future. These new utility values will be used in future recognition processes, and overtime the process distinguishes good from bad instances producing improved decisions. In IBLT, expertise is achieved by the acquisition and refinement of SDU instances. Overtime, in similar contexts, greater number of new SDU instances would be generated by the accumulation and recombination of previous SDU instances. Since instances are evaluated according to previous knowledge and refined based on task feedback, future instances are expected to have higher probability of success.

### 3.5. Summary

IBLT is a theory including five learning mechanisms: instance-based knowledge, recognition-based retrieval, adaptive strategies, necessity, and feedback updates. These mechanisms apply in a decision-making process composed of the steps: recognition, judgment, choice, and feedback. We expect that IBLT will account for the development of decision-making skills in dynamic environments.

First, IBLT provides a realistic account of the use of experience in dynamic, uncertain environments. Knowledge in the form of instances accumulates with practice and instances capture the decision made and the accuracy of the outcome. Second, IBLT proposes a recognition-based use of that instance-based knowledge. The use of previous knowledge depends on the similarity of the current situation to past situations. The similarity of two situations increases with practice on the task, producing the focus of attention to relevant task cues. The third proposal is the adaptive selection of strategies. Heuristics are used if the situation is not similar enough to past experiences. With practice in the same task context, similarity will increase, reducing the use of heuristics and increasing the use of instances overtime. The fourth proposal is a necessary mechanism to make a choice between searching for more alternatives and executing the current best option. This is a combination of optimizing and satisficing strategies in decision making. Finally, the fifth proposal is the knowledge of results to refine and upgrade the SDU's utility slot, resulting in an improved use of experience.

A key characteristic of IBLT is its implementation within the principles of cognition in the ACT-R cognitive architecture (Anderson & Lebiere, 1998). IBLT's mechanisms supported by this cognitive architecture are unique efforts in capturing the instance-based process of learning in a dynamic, real-time decision-making environment. The implementation of IBLT into ACT-R is presented next.

## 4. CogIBLT: an ACT-R implementation of IBLT

CogIBLT is a cognitive model of learning developed from IBLT and framed by the ACT-R 4.0 principles (Anderson, 1993; Anderson & Lebiere, 1998). CogIBLT interacts in real-time with a simulation of a dynamic task. ACT-R is a cognitive modeling architecture built on principles of cognition and supported by a large number of empirical studies in memory, learning and problem solving. In ACT-R, it is possible to represent knowledge in two forms: (1) procedural. If-Then rules or productions, and (2) declarative, chunks. It is also possible to retrieve this knowledge according to a set of performance/learning methods. ACT-R is a hybrid architecture with rules as well as chunk-based learning mechanisms. Chunks encode small, independent patterns of information as sets of slots with associated values. CogIBLT makes use of both forms of knowledge representations, but learning relies on chunks. According to ACT-R there are two sources of chunks: encoding from the environment and the action of a production rule. ACT-R may create a chunk by attending to an object or by solving a goal. This is similar to the proposal from Logan's theory, storing in an instance the results of computations (Logan, 1988). ACT-R, however, avoids creating duplicate chunks, instead it merges chunks and combines their strength. Declarative knowledge is retrieved based on the level of activation associated with a chunk.<sup>2</sup>

### 4.1. Dynamic decision-making task

CogIBLT interacts with a simulation of a complex DDM task called the Water Purification Plant (WPP). WPP is an isomorph of a real world-scheduling task in an organization with large-scale logistical operations, the United States Post Office (Letch, Ballou, & Harter, 1997), but WPP is explainable in less than 1 h and a trial can be completed within few minutes (the experimenter controls the pace of the simulation). A screen shot of the WPP simulation is provided in Fig. 2.<sup>3</sup>

WPP simulates a water distribution system. The system is made up of chains of tanks and each chain is assigned a particular deadline by which a participant must distribute all the water out of the chain. The distribution of water occurs by participant opening or closing pumps in the chain. The simulation is a dynamic environment, where quantities of water in any of tanks may increase (water input from outside of the system) without the participant's knowledge. WPP runs according to a scenario that describes the patterns of water arrival to the system. The scenario defines the time, the amount of water, and the destination tank. WPP is a resource allocation task, where the maximum number of pumps opened is five. The decisions that a participant makes are interconnected: opening a pump in one chain may prevent one to open the water flow in another chain. The main performance indicator in this task is the total number of gallons of water missed. These are the units of water not pumped before their deadlines. After each deadline, the simulation shows the total number of gallons of water missed (see top left corner of Fig. 2). The missed water score is updated after each deadline continuing until the end of the simulation (10:00 p.m.), when the total number of gallons missed is shown on the screen. The best performance in the WPP is 0, indicating all of the water was delivered on time. An optimal solution is defined by a series of activations and deactivations of pumps on the basis of the perceived opportunities. In WPP there might be many possible decision sequences for



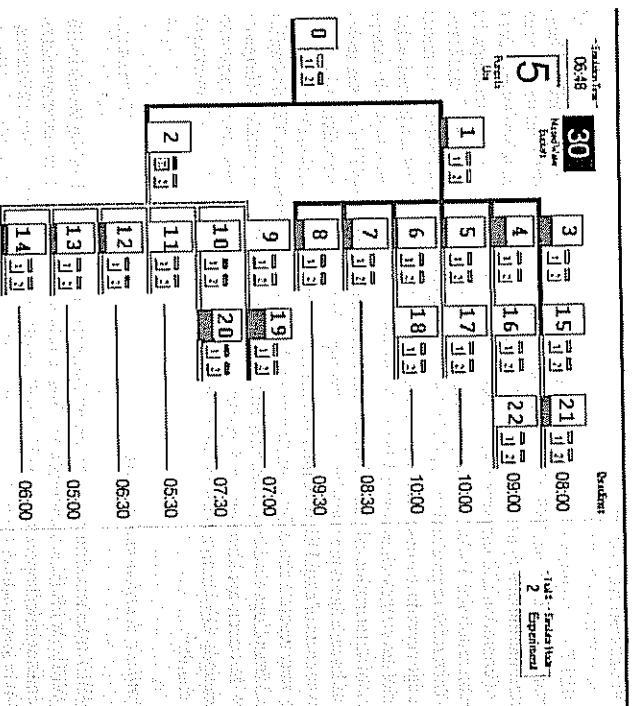


Fig. 2. The layout of the WPP simulation. The simulation time is 6:48 p.m., the operator has missed 30 gal of water, five pumps, the maximum number of pumps, are activated. Water comes from outside of the system and moves continuously on the activated pumps from left to right towards the deadlines. The operator decides when to activate and deactivate pumps while the simulation time is running.

activating and deactivating pumps and achieving the optimal performance. The solution to this task is not unique. We have tested the results of consistent application of different heuristics in WPP.<sup>4</sup> The results show the time heuristic works better than other more complex heuristics. The reason for the success of the time heuristic is in its fit to the structure of the dynamic task. In DDM, decision makers need to relate environmental changes to the timing of their own decisions (Kershoid & Raaijmakers, 1997). We use the time heuristic in CogIBLT to calculate the utility of SDUs in atypical situations.

#### 4.2. Structure of SDUs and the decision process in CogIBLT

We have implemented SDU instances into ACT-R chunk structures. In WPP, a situation is defined by the attributes of a tank, for example, time of evaluation, amount of water, chain value, and deadline. Decisions include activation and deactivation of the pumps associated

Time	Tank	Wateramt	Chain	Utility	Decision
1:38	0	5.0	4	162	Activate

Fig. 3. An example of a SDU for the WPP task. The situation is defined by: time, tank, water amount, and chain. The decision is to activate, and the utility is 162 min.

with a tank. A utility measure, called time heuristic, is the time to reach a deadline. Fig. 3 represents the structure of SDU instances in CogIBLT.

Fig. 4 shows a flowchart of the decision process and the production rules involved in CogIBLT.<sup>5</sup>

Productions are organized in the context of the proposed decision-making process in a goal-centered structure. Verbal protocols,<sup>6</sup> extensive observations of multiple players of WPP, as well as the authors' extensive practice in the task, were the main sources of knowledge to shape the productions and their organization in CogIBLT. Productions are organized according to the type of pursued goal. There are seven productions in CogIBLT. This is a remarkably simple learning model for such complex task as WPP. The complexity of the model relies on the accumulated knowledge in the form of SDUs and their manipulation by both ACT-R and CogIBLT. The productions shown in Fig. 4 are presented next.

Initially, the goal is to meet the first deadline at 5:00 p.m. Production 1 compares the current deadline to the environment and updates SDU's utility if a deadline has been reached. Depending on the environmental conditions, the model may decide to select a tank to evaluate for activation (Production 3) or deactivation (Production 2). The recognition process starts with the selection of tanks. We have not yet implemented the dynamics of attention proposed in IBLT, but we have defined two mechanisms for alternative selection. First, a random function selects tanks out of the possible alternatives in random order. Decision makers are expected to develop focused attention on task-relevant factors. In our verbal protocols, inexperienced users reported random strategies, while more experienced users reported better understanding of deadline, remaining time, amount of water, and chain. A second function selects tanks from the most to the least urgent according to the time remaining to reach a deadline.

Production 6 determines if the situation is typical by comparing the attributes of the selected tank to previous similar situations using ACT-R's partial matching mechanism. Partial Matching (PM) allows the retrieval of a chunk even when it only partially matches a production condition. PM is a useful mechanism when the slots' values of chunks change overtime, since it is unlikely to find an exact match to a particular situation. Production 6 uses PM to compare time, volume of water, and chain values of a current situation to previous SDUs. If there is at least one previous instance similar enough to the current environmental situation, then the situation is classified as typical, otherwise the situation is atypical. If the situation is typical, Production 6 determines the utility from the SDU instances stored in memory. To model the way utilities are combined from past decisions we use ACT-R's Blending mechanism.<sup>8</sup> Blending is a variation of PM that allows the retrieval of an aggregate result set of memory chunks. Blending is described as the value that minimizes the sum of the squared similarities with the values proposed by each chunk, weighted by their probabilities of retrieval. For WPP, Blending retrieves SDUs produced at approximately the same time with about the same amount of water

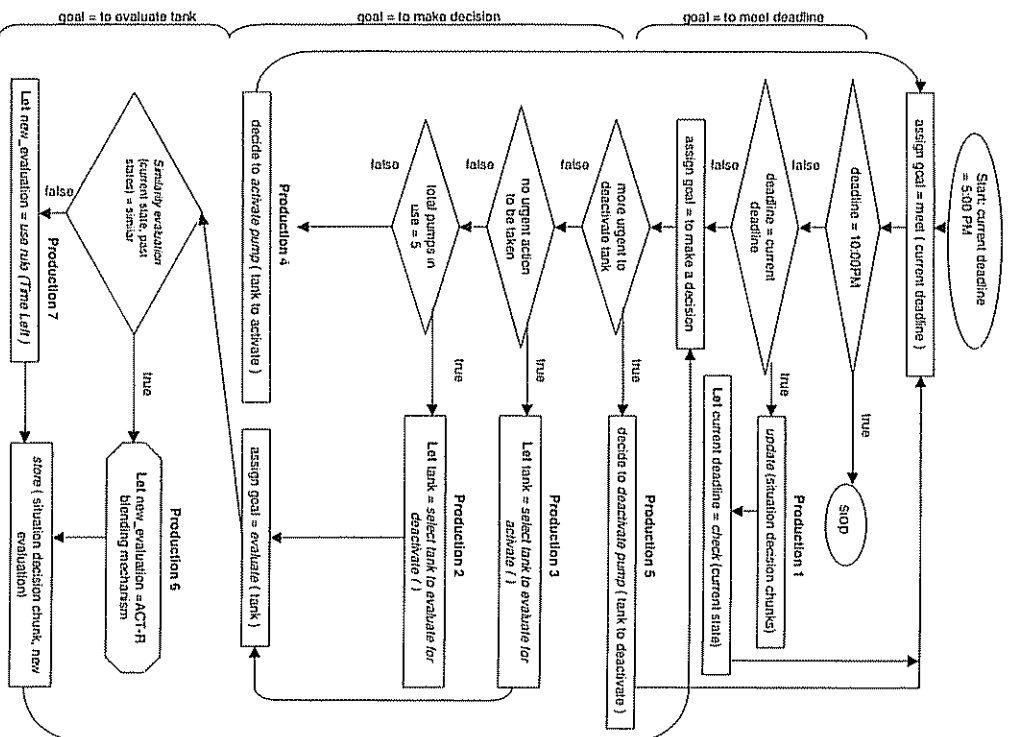


Fig. 4. Flow-chart description of the production rules for CogBLT in the WPP task. Seven productions (numbered 1-7) are organized according to the goal pursued. For example, when the goal is to evaluate a tank, Production 6 or 7 can be used.

Trial	Time	Tank	Water Amount	Chain	Decision	Utility
1	129	9	3.19	1	Activate	306
5	172	15	12.00	1	Activate	309
8	130	9	3.19	0	Activate	293
10	128	9	3.19	1	Activate	297.83

Fig. 5. Blending the utility into a new decision. The current situation in the bottom SDU compares to three SDUs from the past with activation above threshold. The utility of the current situation is calculated by Blending using the utility of the previous three SDUs.

and on a tank with the same chain value as current situation. Fig. 5 shows an example of a situation evaluated in Trial 10 at 2:08 p.m. (minute 128) in Tank 9.

This situation is similar to three previous situations stored in memory. The utility value for the current situation is not the value for one particular instance in the past, but a new aggregated utility value calculated by Blending. Blending pools the strengths of separate chunks (activation) that are similar to the current situation, therefore the new SDU instance has a utility value representing the knowledge of similar instances from the past. Blending is similar to mechanisms in other systems that can combine advice from multiple instances (Hintzman, 1984, 1986) but it features an important difference. The influence of each instance on the consensus decision is weighted by its recency (as reflected in memory decay). Frequency of to the current situation but also of its recency (as reflected in memory decay). Frequency of rehearsal, associative priming and stochastic noise. Thus, Blending also incorporates many cognitive phenomena in a basic mechanism to combine instances. The similarity and the number of previous instances used in Blending may vary from person to person and overtime. In CogBLT, we have implemented a mechanism to manipulate the perception of similarity called similarity-rate. The highest similarity-rate of 1, implies that Blending retrieves *identical* SDU instances to the current situation. In general, with higher the similarity-rate less SDU instances are retrieved and with lower the similarity-rate more SDU instances are retrieved. Therefore, a high similarity-rate relates to low probability of confronting typical situations, while low similarity-rate relates to high probability.

If the situation is atypical, Production 7 calculates the utility of a tank using heuristics. Our cognitive model allows the insertion of any heuristic in the judgment process. But, as previously discussed, the time heuristic is expected to be successful and applicable to other DDM tasks. The time heuristic in WPP recommends decision makers to activate pumps with the closest deadline. The utility of a decision in a particular situation is calculated by subtracting the deadline time of the tank from the current simulation time. For example, if the decision maker intends to activate a pump for the tank with the 5:00 p.m. deadline and the current time is 4:00 p.m., the utility for the current situation is 60 min. For each judgment a new SDU instance is generated. As knowledge accumulates in the form of SDUs, the situations confronted will be more similar to past situations, causing a gradual transition from heuristics to similarity-based retrieval.



The selection of tanks and their evaluation continues unless the necessity level indicates to stop the alternative search. A choice is made to activate (Production 4) or to deactivate (Production 5) a pump. The necessity level in WPP is currently implemented as the time left to a deadline. This is an adjustable threshold beyond which the evaluation of alternatives needs to stop and a decision needs to be made. This threshold may vary from person to person and overtime. For example, we have observed that some decision makers in WPP reduce the number of decisions they make overtime. Apparently, some people learn to wait until the right moment to react to a situation. Human data presented in the results section confirms this observation. The result of the choice step is the best alternative in the evaluation. The execution of actions changes the WPP environment by activating or deactivating a pump and the evaluation of the deadline and other alternatives continues in a cycle.

Production 1 refines the utility value of SDUs according to the performance feedback provided in WPP.<sup>9</sup> Feedback penalization makes those chunks that produced bad performance less likely to be chosen in the future by incrementing the utility value in proportion to the number of gallons missed. All SDU instances produced between the previous and the current deadline are penalized by a percentage of the gallons missed in time units. The percentage to which the feedback updates the utility of chunks is called the learning-rate. This parameter allows variability in the perception of feedback. A learning-rate of 0 indicates that feedback is not used to re-evaluate the decisions made. A learning-rate of 1.0 indicates that the utility of each SDU will be modified by 100% of the value of the missed buckets in time units. For example, regard the SDU shown in Fig. 3. This SDU indicates that at minute 138 or 2:30 p.m., Tank 0 was selected for activation resulting in a utility value of 162.<sup>10</sup> Assuming that at the first deadline of 5:00 p.m., 2 gal of water were missed, the feedback process will update the utility value to 166: 100% of 2 gal  $\times$  2 min pumping-rate. This process of SDU refinement together with the chunk activation makes "bad" chunks less likely to be selected in future actions.

#### 4.3. Summary

CogIBLT is a cognitive model based on IBLT, implemented in ACT-R 4.0, designed to interact with WPP in real-time. CogIBLT uses several parameters from the ACT-R architecture and proposes some other mechanisms to capture human behavior in dynamic environments. Table 1 summarizes the CogIBLT parameters and their use in the model.

There are several instance-based models developed in ACT-R including tasks such as the Sugar Factory, Backgammon, Air-Traffic Control, and estimation of large arithmetic facts (Lebiere, 1999; Lebiere, Anderson, & Botnell, 2002; Lebiere, Wallach, & Niels, 1998; Santer, Anderson, Lebiere, & Lovett, 2000). CogIBLT is, however, the only instance-based learning model that unites several learning mechanisms applied to DDM in addition to the mechanisms provided by ACT-R. CogIBLT is based on a theory of decision making in dynamic environments that perhaps may extend the current functionality of ACT-R. The next section reports some simulation experiments designed to test the propositions from IBLT.

Table 1  
CogIBLT mechanisms and parameters

IBLT step	Mechanism/parameter	Effect
Recognition	CogIBLT mechanisms	Retrieval of alternatives
		Random process Sorted process
		Selects alternatives randomly (tanks in WPP)
Judgment	CogIBLT mechanisms	Evaluation of alternatives
		Heuristic Blending
		Determines the utility value based on a heuristic (in WPP time heuristic) Determines the utility value based on retrieval of similar situations from memory A percentage of linear similarity between two situations
		High (1.0) Low (0.1)
		Only 100% matching chunks are used Any chunk can be used, since no similarity is required
Choice	CogIBLT parameters	Necessity level
		On
		A threshold value to stop the evaluation of alternatives (in WPP is a time limit over which no more tanks are evaluated) Alternatives are evaluated only if the current best utility value has not surpassed the threshold value (in WPP threshold may vary from 1 to 480 min) All alternatives are evaluated before a decision is made
Feedback	CogIBLT parameters	Learning-rate
		Off
		A percentage in which the feedback will influence the utility of a SDU
		High (1.0)
		Feedback effects are accounted 100% over the utility
		Low (0)
		Feedback does not modify the utility of SDUs

#### 5. Simulation experiments

We designed four studies to demonstrate each of the theoretical concepts proposed by IBLT. In addition, these experiments show how CogIBLT was tuned to match human data. Our intention was not to do a sensitivity analysis of ACT-R parameters and learning mechanisms, but to test the IBLT propositions. The order of the computational experiments is the order of the steps of IBLT (Fig. 1). Next, we present the process used to collect human data and results, the IBLT experiments, performance comparisons to human data, and process and individual comparisons.

### 5.1. Human data collection

We recruited 14 students from local universities to run the WPP simulation. All participants ran the simulation 18 times using a standard scenario at 1,008 gal capacity six times per day during 3 consecutive days. Each simulation trial lasted 8 min. The first day participants were given task instructions using the WPP simulation at the slowest pace of 30 min. They were given the task goal and instructions on how to perform the task. While running the simulation, participants practiced activation and deactivation of pumps. We did not allow participants to finish the instructions trial (finishing time is 10:00 p.m.), but rather we stopped the simulation at the first deadline, 5:00 p.m. During the instructions, we made the participants aware of deadlines, simulation time, and the water travel paths. Participants were told that different amounts of water may come from outside of the system to any of the tanks at any time, but they were not given information on the amount of water to process or the time of water arrival. They were instructed to do their best to process all the water that appeared within the system, but no particular strategy was taught.

Fig. 6 shows the participants' average performance for each of the 18 runs of the simulation. A within-subjects analysis of performance for our participants shows significant learning over the 18 trials,  $F(1, 17) = 6.871$ ,  $p < .001$ . Table 2 shows descriptive statistics per trial. On average in the first two trials these participants performed worse than the simulation making random activations of pumps. Also, on average these participants did not reach the score predicted by the time heuristic (time heuristic score is 58 gal).<sup>11</sup>

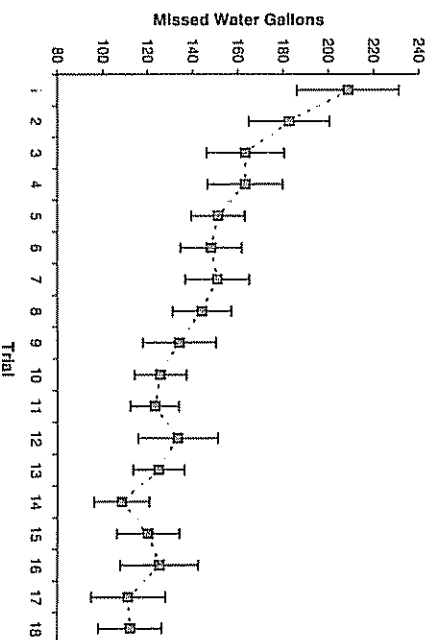


Fig. 6. Participants' average performance for 18 trials. The performance measure is the number of gallons of water missed. The overall mean is 140.55 and the average standard deviation is 20.85.

Table 2  
Human performance per trial: general statistics

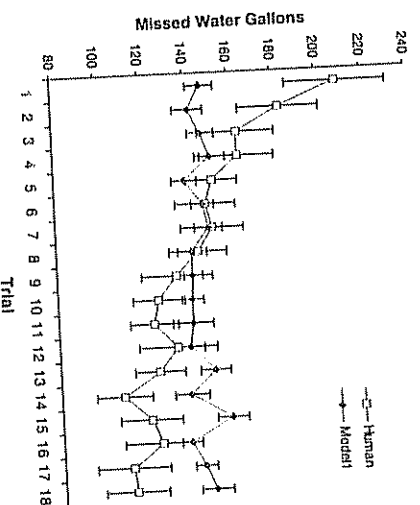
Trial	N	Minimum	Maximum	Mean	SE	SD
1	14	72	389	208.79	22.58	84.49
2	14	97	320	182.71	17.86	66.81
3	14	52	303	163.21	17.27	64.61
4	14	76	360	163.14	16.68	62.43
5	14	87	247	151.21	11.85	44.34
6	14	81	255	148.07	13.50	50.50
7	14	58	218	150.79	14.31	53.54
8	14	87	239	144.07	13.02	48.71
9	14	83	292	134.00	16.24	60.78
10	14	79	240	125.57	11.63	43.51
11	14	73	226	123.07	10.88	40.72
12	14	48	284	133.36	17.68	66.14
13	14	69	203	124.86	11.41	42.70
14	14	42	218	108.43	12.35	46.20
15	14	52	243	120.07	14.01	52.44
16	14	47	243	125.07	17.24	64.50
17	14	22	234	111.29	16.53	61.84
18	14	25	220	112.14	14.09	52.73

### 5.2. Data collection from CogIBLT

Our methodology was to start with a model that has none of the mechanisms and parameters proposed by IBLT. We created simulation experiments to incorporate one-by-one the mechanisms and parameters proposed by IBLT. Our initial state was an ACT-R model with the productions described in Fig. 4 and the chunks encoding the initial system state. The ACT-R default parameters were kept constant throughout these experiments.<sup>12</sup> Each of the models we present in the experiments ran 14 simulated subjects, 18 times per subject using a standard scenario. Each simulation trial ran at a real-time rate of 8 min. The feeding heuristic was the time heuristic. Notice that human data were the baseline for comparison. In each of the experiments, the averaged model's performance was compared to the averaged human data over the 18 trials.<sup>13</sup> We used two measures of goodness of fit as suggested by Schum and Wallach (2002). First, Pearson's  $r^2$  was used<sup>14</sup> to measure how well the relative trend magnitudes were captured by each model. Second, the Root Mean Squared Standard Deviation (RMSSD)<sup>15</sup> measured deviations from exact locations. With each of the experiments, we present a table showing step-by-step calculations per trial. These calculations compare human data to model data, not model to model.

### 5.3. Experiment series 1: recognition process

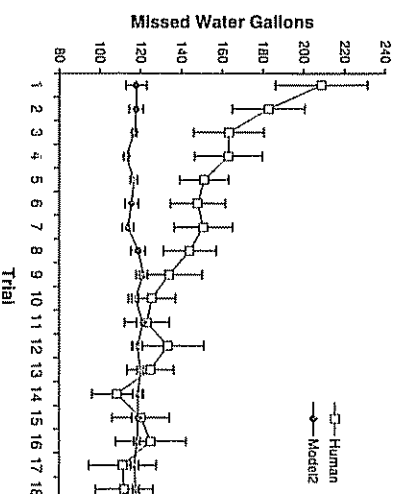
Initially, we ran a model, Model 1, to retrieve alternatives randomly. The model also determined the utility of SDUs based on the time heuristic throughout the trials and did not use Blending. It evaluated all alternatives before a decision was made (no stopping rule) and had zero learning-rate which means feedback had zero effect on the model.



Trial	Human Missed Gallons	Human SE	Model1 Missed Gallons	Model1 SE	$r^2$	MAD	MSAD	PW5SCI	RMSD	RMSSD
1	208.79	22.60	147.43	6.10	0.02	20.42	1.35	61.11	25.96	1.84
2	182.71	17.90	142.07	6.80						
3	163.21	17.30	142.21	6.10						
4	163.14	16.70	150.93	6.90						
5	151.21	11.90	139.07	5.60						
6	148.07	13.50	147.00	5.10						
7	150.79	14.30	149.29	6.20						
8	134.00	13.00	141.36	6.50						
9	125.57	11.60	140.93	4.30						
10	123.07	10.90	140.86	5.00						
11	133.36	17.70	139.29	9.20						
12	124.86	11.40	149.64	6.70						
13	108.43	12.30	138.64	7.50						
14	120.07	14.00	158.86	7.30						
15	125.07	17.20	138.00	4.40						
16	111.29	16.50	143.79	4.90						
17	112.14	14.10	146.36	7.10						
18										

Fig. 7. Model 1. Random search of alternatives, utility determined by the time heuristic, no Blending, no necessarily value, and no feedback mechanisms.

Fig. 7 shows the results from this model compared to human data. Because of the absence of learning, Model 1 did not capture the relative direction effects of human data ( $r^2 = .02$ ). Average human performance started significantly worse than Model 1 but became significantly better by the final trials (RMSD = 1.64).



Trial	Human Missed Gallons	Human SE	Model2 Missed Gallons	Model2 SE	$r^2$	MAD	MSAD	PW5SCI	RMSD	RMSSD
1	208.79	22.60	117.86	5.00	-0.12	25.19	1.56	61.11	35.02	2.02
2	182.71	17.90	117.79	3.40						
3	163.21	17.30	116.93	1.10						
4	163.14	16.70	113.00	1.20						
5	151.21	11.90	118.71	1.70						
6	148.07	13.50	115.57	3.20						
7	150.79	14.30	113.93	2.70						
8	144.07	13.00	118.86	3.50						
9	134.00	16.20	121.50	1.90						
10	125.57	11.60	117.38	1.70						
11	123.07	10.90	121.21	3.10						
12	133.36	17.70	118.57	3.10						
13	124.86	11.40	119.93	1.60						
14	108.43	12.30	118.79	2.50						
15	120.07	14.00	118.79	3.00						
16	125.07	17.20	118.21	1.50						
17	111.29	16.50	117.07	1.90						
18	112.14	14.10	117.71	1.60						

Fig. 8. Model 2. Sorted search of alternatives, utility determined by the time heuristic, no Blending, no necessarily value, and no feedback mechanisms.

To test the effect of a sorted process, Model 1 was modified to select tanks from the environment based on the priority determined by the time heuristic over the course of the 18 trials.

Fig. 8 shows the results from this model. Models 1 and 2 together suggest that humans behave similar to the random alternative selection in the first trials and similar to the sorted alternative selection in the last trials. Since the first two trials show the highest RMSD values

with an average = 14.67, we modified Model 2 and ran the first two trials under a random selection of alternatives.

The results from this model, Model 3, are shown in Fig. 9. The relative trend measure improved,  $r^2 = .46$ , indicating a closer trend to human data. Still, the approximation to human data as measured by RMSSD indicates the model was far from capturing the human's exact performance values, especially during the first trials of the experiment. Model 3 evaluated alternatives based exclusively on the time heuristic. Although Model 3 generated SDUs for every evaluation it performed, evaluations were performed by time heuristic only rather than based on examples from memory. The next step is to demonstrate the effects of Blending.

#### 5.4. Experiment series 2: judgment

In this experiment series, we tested the effects of Blending and similarity-rate. When past examples are used to evaluate the utility of an alternative, the similarity-rate determines the examples used from the past according to how well they match the current situation. The highest similarity-rate, 1.0, retrieves past examples only if they match perfectly the alternative evaluated. Similarity-rate determines the percentage of linear similarity between two tank situations. Starting with Model 3, we let CogIBLT use Production 6 (Fig. 4) to activate Blending. If similar examples to the current situation exist, then CogIBLT retrieves all past similar SDUs and blends them to produce the utility of the current situation. Otherwise, the model evaluates the current alternative using the time heuristic. Model 4 used Blending, with highest similarity-rate, 1.0. Production 6 activated only if the exact same situation happened in the past. Again, in DDM it's rare to find two identical decision situations. We expected that Model 4 would match human data in a similar way as Model 3. The results from Model 4 are shown in Fig. 10.

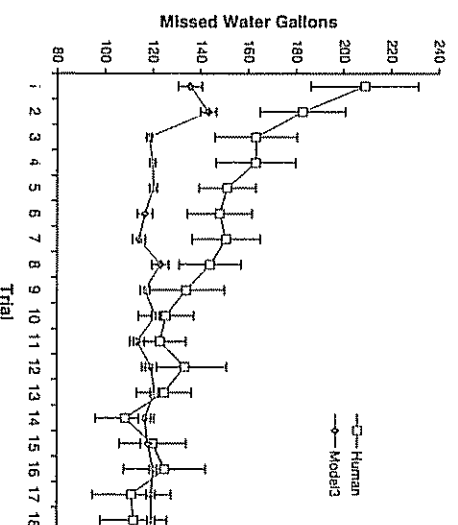
Model 4 captured the relative trend of human data more closely than Model 3,  $r^2 = .64$ . The addition of Blending with the similarity-rate of 1.0, made model's performance much closer to human performance in the first two trials. Model 4's performance was better than the human data performance in Trials 3-8, as indicated by the RMSSD.

Next, CogIBLT's behavior was tested with the lowest similarity-rate, 0.01. This value indicates that CogIBLT would blend SDUs from the past if they fit the current situation by at least 1%. We expected that many past situations were similar to the current alternative in at least 1% because we were using the same scenario throughout the 18 trials in the same task.

Fig. 11 shows the results from this model, Model 5. The effects of Blending with similarity-rate of 0.01 are considerable as Model 5 captures the trend of human data quite closely with  $r^2 = .81$ . Model 5 also predicts the exact performance within the 95% confidence interval in most of the trials, average RMSSD = 1.71. Blending with the lowest similarity-rate made the model's performance worse than the humans' performance for the middle portion of the trials.

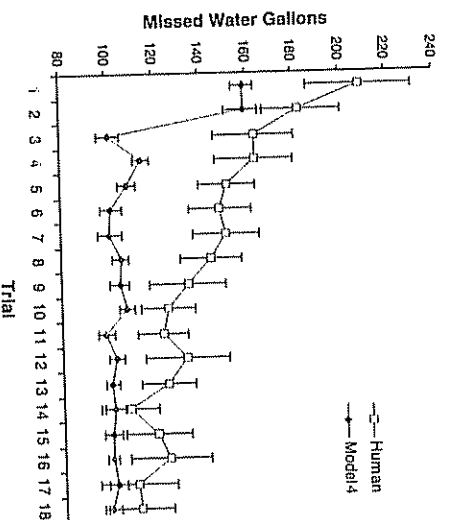
#### 5.5. Experiment series 3: choice

Models in all previous experiments did not make use of a stopping rule. This means previous models evaluated all alternatives before executing an action, unless there was an external event during the evaluation process, for instance, incoming water. Here we explore the effects



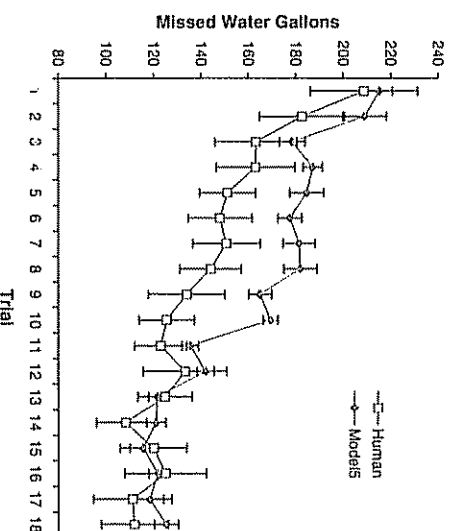
Trial	Human Missed Gallons	HumanSE	Model3 Missed Gallons	Model3 SE	Model3 abs(data-model)	Model3 abs(data-model) SE	MSAD	PWSOI	RMSD	RMSSD
1	208.79	22.60	135.57	5.00	73.21	3.24	1.40	61.11	29.18	1.73
2	182.71	17.90	143.29	3.40	39.43	2.20	0	0	1554.61	4.85
3	163.21	17.30	118.43	1.10	44.79	2.59	0	0	2005.76	6.70
4	163.14	16.70	119.93	1.20	43.21	2.59	0	0	1867.47	6.70
5	151.21	11.90	120.21	1.70	31.00	2.61	0	0	961.00	6.79
6	148.07	13.50	116.79	3.20	31.29	2.32	0	0	978.80	5.37
7	150.79	14.30	114.29	2.70	36.50	2.55	0	0	1332.25	6.51
8	144.07	13.00	123.29	3.50	20.79	1.60	1	1	492.05	2.56
9	134.00	16.20	116.79	1.90	17.21	1.06	1	1	286.33	1.13
10	125.57	11.60	121.21	1.70	4.36	0.38	1	1	18.98	0.14
11	123.07	10.90	113.57	3.10	9.50	0.87	1	1	90.25	0.76
12	133.36	17.70	113.36	2.20	14.00	0.79	1	1	196.00	0.63
13	124.86	11.40	120.79	1.60	4.07	0.56	1	1	16.55	0.13
14	106.43	12.30	116.79	2.50	8.36	0.68	1	1	69.84	0.46
15	120.07	14.00	116.00	3.00	2.07	0.15	1	1	4.29	0.02
16	125.07	17.20	120.29	1.50	4.79	0.28	1	1	22.90	0.24
17	111.29	16.50	119.43	1.90	8.14	0.49	1	1	66.31	0.24
18	112.14	14.10	119.50	1.60	7.36	0.52	1	1	54.13	0.27

Fig. 9. Model 3. Random search of alternatives in the first two trials, sorted search after Trial 3. Utility determined by the time heuristic, no Blending, no necessary value, and no feedback mechanisms.



Trial	Human Missed Gallons	Human SE	Model 4 Missed Gallons	Model 4 SE	abs(data - model) data SE	is MSAD < 1.95	(data - model)/2	(data - model)/data SE/2
1	208.79	22.60	158.71	4.50	50.07	0	2507.15	4.51
2	182.71	17.30	158.86	8.10	23.85	1	569.16	1.78
3	163.21	17.30	100.64	5.00	62.57	0	3915.18	13.08
4	163.14	16.70	114.79	3.60	48.36	0	2338.41	8.38
5	151.21	11.90	108.36	3.70	42.86	0	1836.73	12.87
6	148.07	13.50	101.64	4.70	46.43	0	2155.61	11.83
7	150.79	14.30	101.07	5.30	49.71	0	2471.51	12.09
8	144.00	13.00	105.64	3.50	38.43	0	1476.76	8.74
9	134.00	16.20	105.07	4.10	28.93	1	836.86	3.19
10	125.57	11.60	108.07	3.40	17.50	1	306.25	2.28
11	123.07	10.90	98.79	3.50	24.29	0	589.80	4.96
12	133.36	17.70	103.21	3.30	30.14	1	908.58	2.90
13	124.86	11.40	101.07	3.20	23.79	0	565.76	4.35
14	108.43	12.30	101.93	4.50	6.50	1	42.25	0.28
15	120.07	14.00	100.86	3.70	19.21	1	369.19	1.88
16	125.07	17.20	100.57	2.50	24.50	1	600.25	2.03
17	111.29	16.50	102.64	3.90	8.64	1	74.70	0.27
18	112.14	14.10	99.93	3.60	12.21	1	149.19	0.75

Fig. 10. Model 4. Random search of alternatives in the first two trials, sorted search after Trial 3. Utility determined Blending with one similarity-rat, no necessary value, and no feedback mechanisms.



Trial	Human Missed Gallons	Human SE	Model 5 Missed Gallons	Model 5 SE	abs(data - model) data SE	is MSAD < 1.95	(data - model)/2	(data - model)/data SE/2
1	208.79	22.60	215.21	5.50	6.43	0.28	41.33	0.08
2	182.71	17.30	209.07	9.20	26.36	1	694.70	2.17
3	163.21	17.30	178.57	5.40	15.36	0.89	235.84	0.79
4	163.14	16.70	167.21	4.00	24.07	1	579.43	2.08
5	151.21	11.90	184.64	7.10	33.43	0	1177.47	7.89
6	148.07	13.50	177.64	5.00	29.57	0	874.47	4.80
7	150.79	14.30	181.50	6.80	30.71	0	943.37	4.61
8	144.00	13.00	182.00	7.00	37.93	2.92	1438.58	8.51
9	134.00	16.20	164.93	4.80	30.93	1	956.58	3.64
10	125.57	11.60	169.43	3.10	43.86	3.78	1923.45	14.29
11	123.07	10.90	135.50	3.50	12.43	1	154.47	1.30
12	133.36	17.70	142.00	3.50	8.64	0.49	74.70	0.24
13	124.86	11.40	121.64	3.80	3.21	0.28	10.33	0.08
14	108.43	12.30	121.14	4.10	12.71	1.03	161.65	1.07
15	120.07	14.00	115.64	5.40	4.43	0.32	19.61	0.10
16	125.07	17.20	121.86	3.80	3.21	0.19	10.33	0.03
17	111.29	16.50	118.43	5.70	7.14	0.43	51.02	0.19
18	112.14	14.10	125.43	5.00	13.29	0.94	176.51	0.89

Fig. 11. Model 5. Random search of alternatives in the first two trials, sorted search after Trial 3. Utility determined Blending with 0.01 similarity-rat, no necessary value, and no feedback mechanisms.

of the necessary level in the choice step. The necessary levels are now varied as: high = 8 h, medium = 1 h, and low = 1 min. High threshold indicates that any utility value lower than 8 h is urgent. CogIBLT will stop the judgement process and the selection of new alternatives to execute the most urgent action. Medium threshold stops the judgement process if the utility

value of the most urgent option is below 1 h. Finally, low threshold indicates that any utility value below 1 min will stop the judgment process. A high necessity level may classify all the SDUs as urgent, because the remaining time will be less than the total length of the simulation. Intuitively, we might think that in a dynamic environment it is important to react as fast as possible to a situation. On the other hand, we might also expect that stopping the judgment process might degrade performance because not enough alternatives are explored in a decision space. For low necessity levels, more alternatives might be considered but performance may be degraded as CogIBLT may wait too long to take action in an urgent situation.

Starting with Model 5, we varied the necessity level and compared it to human data. We kept the same necessity value throughout the 18 trials.

The results from this manipulation are presented in Figs. 12-14, with the respective tables comparing to humans' learning curves. The results indicate that the inclusion of the necessity level produced a slight improvement in the relative trend as compared to human data. The  $r^2$  measures for all the three models are higher than the one for Model 5 (Model 6: Model 6-High  $r^2 = .86$ , Model 6-Medium  $r^2 = .85$ , Model 6-Low  $r^2 = .90$ ). The average RMSSD indicates that high necessity levels produce the lowest fit to specific performance points (average RMSSD = 2.62) compared to medium and low necessity levels (medium average RMSSD = 1.43, low average RMSSD = 1.26).

### 5.6. Experiment series 4: feedback

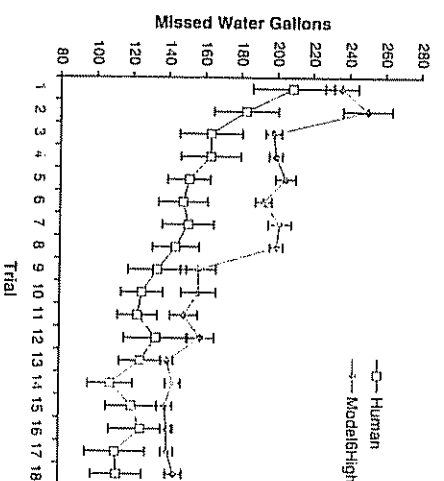
All models used up to this point had unactivated feedback (0 learning-rate parameter). The models learned without considering the effect of the results. Here we test the effects of feedback by updating the learning parameter from 0 to 1.0 in Model 6-Low (the best predictor of humans data). The same value of 1.0 was maintained throughout the 18 trials and, with the parameter set to 1.0, the utility of the SDUs created between the current and the previous deadline will be increased by 100% of the gallons missed after the deadline. Fig. 15 shows the results from Model 7 with learning-rate 1.0.

This model shows better performance compared to other models, but decreases the match to human data. The trend of the learning curve reduces to  $r^2 = .78$  and the average RMSSD increases to 1.57.

### 5.7. Summary of experiments

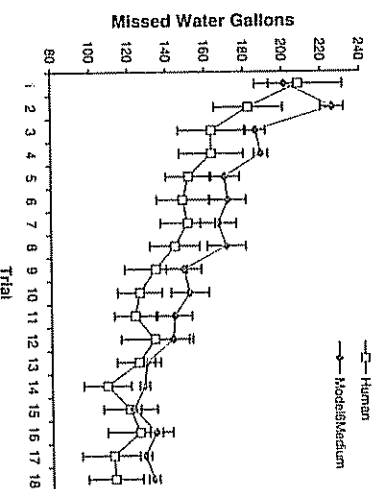
Table 3 summarizes the conditions and the values of comparison to human data.

The base model is Model 1 with random alternative selection, no similarity judgments, no stopping rule and no feedback. The cells in bold indicate the experimental variable modified from the previous model. The  $r^2$  values show how each of the models captures the relative trend magnitudes of human data. RMSSD values indicate how the each of the models captures exact value points of human data. The model that best fits the learning trend of humans is Model 6-Low ( $r^2 = .90$ ). This model is, however, not the one that gives best performance. For example, the minimum average number of gallons missed in Model 6-Low is 114.86 and occurs in Trial 18, while Model 7 shows a minimum average of 97.93 gal of water in Trial 12. Each



Trial	Human		Model 6-High		Model 6-Medium		Model 6-Low		Model 7	
	Missed Gallons	HumanSE	Missed Gallons	Model6-High GH SE	Missed Gallons	Model6-Medium GH SE	Missed Gallons	Model6-Low GH SE	Missed Gallons	Model7 GH SE
1	208.79	22.60	233.89	9.17	250.00	13.74	267.29	3.76	0	4527.37
2	182.71	17.30	197.92	4.46	34.71	2.01	1204.70	14.13	0	1204.70
3	163.21	16.70	199.23	3.71	36.09	2.16	1302.34	4.67	0	1302.34
4	151.21	11.90	204.77	5.64	53.55	4.50	2868.13	20.25	0	2868.13
5	148.07	13.50	192.46	4.55	44.39	3.29	1970.48	10.81	0	1970.48
6	150.79	14.30	201.54	6.48	50.75	3.55	2575.64	12.60	0	2575.64
7	144.07	13.00	199.77	3.69	55.70	4.28	3102.25	18.36	0	3102.25
8	134.00	16.20	156.77	9.92	22.77	1.41	518.44	1.98	0	518.44
9	125.57	11.60	156.92	9.64	31.35	2.70	682.93	7.30	0	682.93
10	123.07	10.90	148.77	7.54	25.70	2.36	680.38	5.56	0	680.38
11	133.35	17.70	158.00	7.71	24.64	1.39	607.27	1.94	0	607.27
12	124.86	11.40	139.85	3.45	14.99	1.31	224.67	1.73	0	224.67
13	108.43	12.30	142.85	4.17	34.42	2.80	1184.57	7.83	0	1184.57
14	120.07	14.00	138.62	4.29	18.54	1.32	343.88	1.75	0	343.88
15	125.07	17.20	139.85	3.35	14.77	0.85	218.29	0.74	0	218.29
16	111.29	16.50	140.08	3.35	28.79	1.74	828.93	3.04	0	828.93
17	112.14	14.10	143.77	4.53	31.63	2.24	1000.23	5.03	0	1000.23
18	112.14	14.10	143.77	4.53	31.63	2.24	1000.23	5.03	0	1000.23

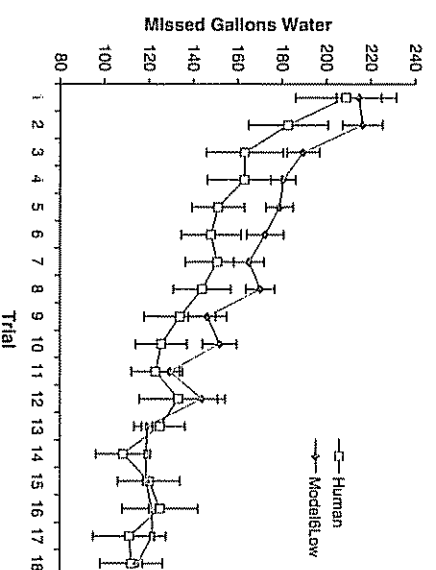
Fig. 12. Model 6-High. Random search of alternatives in the first two trials, sorted search after Trial 3. Utility determined blending with 0.01 similarity-rate, 8 h necessity level, and no feedback mechanisms.



Model 7.									
	$r^2$	MAD	MSAD	Pw50CI	RMSD	RMSSD			
	0.85	18.10	1.26	83.33	20.50	1.43			
Trial	Human Missed Gallons	Human SE	Model6-Medium Missed Gallons	Model6-Medium SE	abs(data - model)	abs(data - model) SE	Is MSAD	(data - model) <sup>2</sup>	(data - model) <sup>2</sup> SE <sup>2</sup>
1	208.79	22.60	201.57	8.14	7.21	0.32	1	52.05	0.10
2	182.71	17.90	226.00	5.99	43.29	2.42	0	1873.65	5.85
3	163.21	17.30	186.36	4.98	23.14	1.34	1	535.59	1.79
4	163.14	16.70	189.07	3.55	25.93	1.55	1	672.29	2.41
5	151.21	11.90	169.92	7.80	18.71	1.57	1	350.02	2.47
6	148.07	13.50	171.62	9.25	23.54	1.74	1	554.92	3.04
7	150.79	14.30	166.71	9.35	15.93	1.11	1	253.72	1.24
8	144.07	13.00	170.93	10.06	26.86	2.07	0	721.31	4.27
9	134.00	16.20	148.57	9.05	14.57	0.90	1	212.33	0.81
10	125.57	11.60	151.71	9.78	26.14	2.25	0	683.45	5.08
11	123.07	10.90	143.79	8.91	20.71	1.90	1	429.08	3.61
12	133.36	11.40	142.86	10.14	9.50	0.54	1	90.25	0.29
13	124.86	11.40	128.93	4.18	4.07	0.36	1	16.58	0.13
14	108.43	12.30	127.71	2.57	19.29	1.57	1	371.94	2.46
15	120.07	14.00	122.00	3.53	1.93	0.14	1	3.72	0.02
16	125.07	17.20	133.50	3.35	8.43	0.49	1	71.04	0.24
17	111.29	16.50	127.79	3.03	16.50	1.00	1	272.25	1.00
18	112.14	14.10	132.14	2.83	20.00	1.42	1	400.00	2.01

Fig. 13. Model 6-Medium. Random search of alternatives in the first two trials, sorted search after Trial 3. Utility determined Blending with 0.01 similarity-rate, 1 h necessity level, and no feedback mechanisms.

of the experimental variables adds incrementally to the full model. Model 7, containing all the learning mechanisms proposed in IBLT. The last column in Table 3,  $r^2$  model comparison, shows how each of the models capture the relative trend magnitudes of Model 7. If we had to choose for a restricted model that best fits the full model, the choice would be Model 6-Medium.



Model 6.									
	$r^2$	MAD	MSAD	Pw50CI	RMSD	RMSSD			
	0.90	14.77	1.03	83.33	17.78	1.26			
Trial	Human Missed Gallons	Human SE	Model6-Low Missed Gallons	Model6-Low SE	abs(data - model)	abs(data - model) SE	Is MSAD	(data - model) <sup>2</sup>	(data - model) <sup>2</sup> SE <sup>2</sup>
1	208.79	22.60	214.64	10.12	5.86	0.26	1	34.31	0.07
2	182.71	17.30	216.36	8.94	33.64	1.88	1	1131.84	3.53
3	163.21	17.30	189.57	7.30	26.36	1.52	1	694.70	2.32
4	163.14	16.70	180.64	5.61	17.50	1.05	1	306.25	1.10
5	151.21	11.90	178.86	6.08	27.64	2.32	0	764.13	5.40
6	148.07	13.50	172.43	8.31	24.36	1.80	1	593.27	3.26
7	150.79	14.30	165.07	6.83	14.29	1.00	1	204.08	1.00
8	144.07	13.00	170.21	6.43	26.14	2.01	0	683.45	4.04
9	134.00	16.20	146.43	8.61	12.43	0.77	1	154.47	0.59
10	125.57	11.60	152.00	7.73	26.43	2.28	0	698.47	5.19
11	123.07	10.90	129.50	5.63	6.43	0.59	1	41.33	0.35
12	133.36	11.40	144.07	10.36	10.71	0.61	1	114.00	0.37
13	124.86	11.40	119.21	2.32	5.64	0.49	1	31.84	0.25
14	108.43	12.30	118.93	0.81	10.50	0.85	1	110.25	0.73
15	120.07	14.00	118.71	1.32	1.36	0.10	1	1.84	0.01
16	125.07	17.20	121.50	1.50	3.57	0.21	1	12.76	0.04
17	111.29	16.50	121.64	1.03	10.36	0.63	1	107.27	0.39
18	112.14	14.10	114.86	2.44	2.71	0.19	1	7.37	0.04

Fig. 14. Model 6-Low. Random search of alternatives in the first two trials, sorted search after Trial 3. Utility determined Blending with 0.01 similarity-rate, 1 min necessity level, and no feedback mechanisms.





### 6.1. Average fit to decision rules

The fit to a rule is a measure of how close SDUs are to what a heuristic prescribes. For each activation decision made by a participant in each trial of the WPP task, we calculated the fit to the time heuristic. The same average fit measure was calculated using the SDUs of decisions made by Model 6-Low since this model gave the closest fit to human data. The following analysis ignores pump deactivation.

Fit values are calculated by comparing the alternative selected by a participant or by the model to the alternative prescribed by the time heuristic:

$$\text{Fit} = 1 - \frac{\text{actual decision} - \text{worst decision}}{\text{best decision} - \text{worst decision}}$$

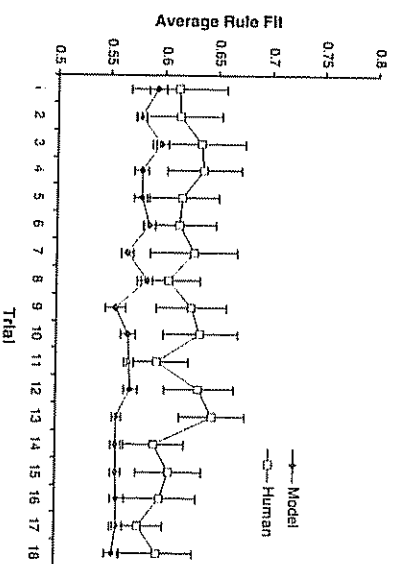
The actual decision is the decision made by the participant or the model. The worst and best decisions are obtained from the alternatives available at the moment of the evaluation, based on the time heuristic. A fit value of zero means the decision was the worst decision that could be made according to the heuristic, while a fit value of one means that the decision is exactly the same prescribed by the time heuristic. We calculated the average fit values for all decisions in each trial. This is similar to the measure used by Payne et al. (1993) for calculating relative accuracy in their effort-accuracy framework for static decisions. In our task, however, the values of the decisions (current decision, best and worst decisions) are dynamic depending upon the specific situation of the system when the decision was made. In Payne et al. studies all accuracy measures are static because the task consists of selecting among previously defined gambles.

Fig. 16 shows the rule fit results for both humans and Model 6-Low. The model's decisions fit the time heuristic less than the human's decisions, but the average of the absolute values of rule fit are close to those from the human's average,  $\text{RMSSD} = 1.65$ . The average fit to the rule shows how closely the decisions fit the time heuristic; however, this measure cannot tell us if judgments were in fact performed by the time heuristic. Although we cannot have this information from humans,<sup>16</sup> we know this data from CogIBLT.

Fig. 17 shows the average percentage of evaluations performed with the time heuristic in each of the 18 trials. In the first trial, 100% of the judgments are performed by the time heuristic (there are no prior examples). Still, in the first few trials a high percentage of judgments are done using the time heuristic. Note that this does not imply a perfect rule fit (see Fig. 16) because not all alternatives are considered. Overtime, CogIBLT increases the reliance on past examples and reduces the use of the time heuristic. By the end of the experiment, in the 18th trial, about 34% of the judgments were made by the time heuristic (about 66% by example). We would expect that humans follow the time heuristic in a pattern similar to that of Fig. 17.

### 6.2. Instance similarity

Instance similarity is a measure of how closely a situation resembles the circumstances from past decisions. All the activation decisions in a trial were sorted by tank and then by the time at which the decision was made. This produced a sequence of activation decisions overtime in each tank and per trial. Decisions in trial  $i$  were compared to decisions in trial  $i - 1$  in their sequential order according to the time at which the decision was made. Similarity was



C. Gonzalez et al. / Cognitive Science 27 (2003) 591-635

Trial	Human	HumanSE	Model	ModelSE	abs (data - model)	abs(data - model) data SE	1a MSAD < 1.98	(data - model)/2	(data - model)/abs SE/2
1	0.61	0.05	0.59	0.01	0.02	0.44	1	0.00	0.19
2	0.61	0.04	0.58	0.01	0.04	0.92	1	0.00	0.84
3	0.63	0.04	0.60	0.01	0.04	0.90	1	0.00	0.81
4	0.64	0.03	0.58	0.01	0.06	1.68	1	0.00	2.81
5	0.62	0.03	0.58	0.01	0.04	1.12	1	0.00	1.25
6	0.61	0.03	0.59	0.01	0.03	0.86	1	0.00	0.74
7	0.63	0.04	0.57	0.01	0.06	1.52	1	0.00	2.32
8	0.60	0.03	0.58	0.01	0.02	0.70	1	0.00	0.50
9	0.63	0.03	0.56	0.01	0.07	2.13	1	0.00	4.55
10	0.63	0.03	0.57	0.01	0.03	0.89	1	0.00	3.82
11	0.59	0.03	0.57	0.01	0.06	1.98	0	0.01	3.91
12	0.63	0.03	0.57	0.01	0.03	0.89	1	0.00	3.82
13	0.65	0.03	0.56	0.00	0.09	2.92	0	0.01	8.54
14	0.59	0.03	0.56	0.00	0.04	1.25	1	0.00	1.58
15	0.61	0.03	0.55	0.00	0.05	1.62	1	0.00	2.64
16	0.50	0.03	0.56	0.00	0.04	1.20	1	0.00	1.44
17	0.58	0.02	0.56	0.00	0.02	0.85	1	0.00	0.74
18	0.60	0.03	0.55	0.00	0.04	1.21	1	0.00	1.47

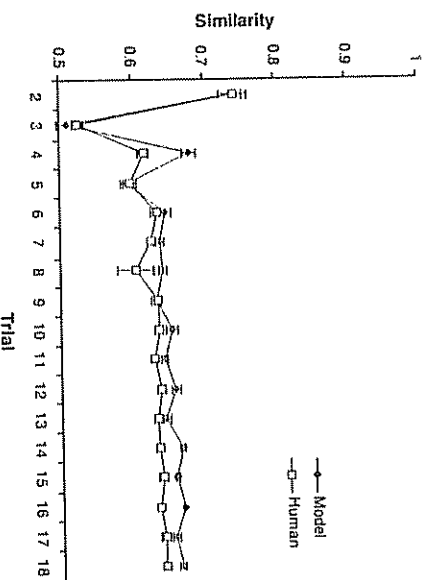
Fig. 16. Average rule fit, comparison between human data and Model 6-Low.

calculated based upon two variables: the time of the decision and the amount of water in the tank at the time the decision was made (water was converted to time units). The following formula was used:

$$\text{Similarity} = 1 - |\alpha(\text{decision}_i - \text{decision}_{i-1}) + (1 - \alpha)(\text{accumulated similarity}_{i-1})|$$

This calculation compares time and water values to those of a previously made decision in prior trials and considers the accumulated similarity from previous trials. The  $\alpha$  value gives different weights to the similarity of the most recent decisions compared to all previous ones.<sup>17</sup>





Trial	r2		MAD		MSAD		PW95CI		RMSD		RMSSD	
	Model	SE	Model	SE	Model	SE	Model	SE	Model	SE	Model	SE
Total	0.74	0.02	0.74	0.02	0.02	0.00	0.19	0.00	0.00	0.00	0.04	0.04
1	0.53	0.01	0.51	0.01	0.02	0.00	1.92	0.00	0.00	0.00	3.70	0.00
2	0.62	0.01	0.68	0.01	0.06	0.06	7.53	0.00	0.00	0.00	56.64	0.00
3	0.60	0.01	0.59	0.01	0.00	0.00	0.49	0.00	0.00	0.00	0.24	0.00
4	0.64	0.01	0.65	0.01	0.01	0.01	1.13	0.00	0.00	0.00	1.27	0.00
5	0.63	0.02	0.64	0.01	0.01	0.01	2.87	0.00	0.00	0.00	8.24	0.00
6	0.61	0.01	0.63	0.01	0.04	0.04	1.47	0.00	0.00	0.00	2.17	0.00
7	0.63	0.01	0.63	0.01	0.00	0.00	0.30	0.00	0.00	0.00	0.09	0.00
8	0.63	0.01	0.65	0.01	0.02	0.02	3.10	0.00	0.00	0.00	9.62	0.00
9	0.63	0.01	0.64	0.01	0.01	0.01	2.74	0.00	0.00	0.00	7.52	0.00
10	0.63	0.01	0.64	0.01	0.02	0.02	4.24	0.00	0.00	0.00	18.01	0.00
11	0.63	0.01	0.64	0.01	0.01	0.01	1.65	0.00	0.00	0.00	2.75	0.00
12	0.63	0.01	0.64	0.01	0.01	0.01	5.82	0.00	0.00	0.00	33.91	0.00
13	0.64	0.01	0.67	0.00	0.03	0.03	4.24	0.00	0.00	0.00	10.48	0.00
14	0.64	0.01	0.67	0.00	0.03	0.03	7.80	0.00	0.00	0.00	60.78	0.00
15	0.64	0.01	0.66	0.01	0.01	0.01	4.49	0.00	0.00	0.00	5.14	0.00
16	0.64	0.01	0.66	0.01	0.02	0.02	4.49	0.00	0.00	0.00	20.19	0.00
17	0.64	0.01	0.66	0.01	0.02	0.02	4.49	0.00	0.00	0.00	20.19	0.00
18	0.64	0.01	0.66	0.01	0.02	0.02	4.49	0.00	0.00	0.00	20.19	0.00

Fig. 19. Average similarity with  $\alpha = 0.5$ .

Results from the first series of experiments suggest that, experienced decision makers may select alternatives according to the relevance of task cues to their decision accuracy while inexperienced decision makers follow a random alternative selection process. Both model and human data show that overtime SDU instances are increasingly similar to the most recent and past decisions in terms of environmental conditions.<sup>18</sup> Similarly to the chunking theory and the instance-based recognition model, in IBLT skills develop through the accumulation of instances

and the recognition of features similar to the probe (Hintzman, 1984; Simon & Guber, 1996). However, many issues remain to be investigated with respect to recognition in DDM. First, our approach involved discrete tests of parameters fixed throughout 18 trials. IBLT, however, suggests a gradual upgrade of the parameters overtime. More empirical data is needed to understand how humans may adapt these mechanisms as they practice in a dynamic task. In most situations, these mechanisms should be modeled as gradual functions. For example, we do not expect humans to suddenly change their attention patterns from random to sorted. We believe that in DDM focused attention develops gradually, with practice in the task, as suggested in psychology (Haider & Frensch, 1996; Logan, 1988). This gradual transition of the model's parameters was not implemented in the current CogIBLT. Second, the role of similarity in decision making needs to be studied further. The importance of similarity in decision making was highlighted by Tversky (1977). But since then, only a few studies have addressed this topic in the behavioral decision-making literature (Leland, 2000; Rubinstein, 1988). Psychological research has identified the relationships between decision-making and high level cognitive processes such as similarity and categorization (Markman & Medin, 1995; Medin, Goldstone, & Markman, 1995) but their connections to dynamic and complex tasks is not clear. We believe that, in DDM, similarity is essential to learning by practicing interactively with a dynamic environment. Similarity influences the recognition process, and the use of heuristics and instances in judging the utility of choices.

Results from the second experimental series indicate a reduction in the use of heuristics and increase of memory retrieval overtime, supporting Logan's (1988) propositions. However, very little is known about the adaptation of judgment strategies and their relationship to memory retrieval. IBLT suggests judgment in DDM is of two flavors: heuristic-based and instance-based. Based on the similarity of the situations to which the decision maker is exposed, judgment turns from heuristic-based to instance-based. In the current study, we used the same pattern of exogenous events throughout the 18 trials. Having the same exogenous events at the same times increases the probability of finding similar current situations to past experiences. IBLT predicts heuristic-based judgments in situations with more diverse environmental events, but it also predicts a switch towards instance-based judgments as more interactive practice is acquired within the same task context. In this study, we used the heuristic for our models. We believe that time heuristic is universal to DDM, but other heuristics may also lead to good performance. Research in the use of heuristics for decision making is taking an interesting turn since the proposal of the "adaptive toolbox" by Gigerenzer and Todd (1999). The study of their proposed heuristics in the context of DDM is a challenge worth pursuing. Also, results from this experimental series provide support to the concept of Blending. IBLT suggests decision makers "blend" their past knowledge to come up with the utility of a situation-decision condition, rather than using one specific example from the past. A similarity-rate of 1 (retrieving examples identical to past instances) resulted in almost immediate improvement of performance but not much learning overtime. Thus, the model's learning curve was distant from humans' learning curves. Retrieval of past solution to the same problem may be an effective problem solving strategy.<sup>19</sup> The results suggest this is not what humans do in dynamic environments. This may be because in dynamic situations it is rare that the same exact problem will occur more than once. The introduction of Blending made model performance closer to human performance. This similarity-based memory

retrieval is comparable to many other learning theories (Hintzman, 1984; Medin & Schaffer, 1978; Medin et al., 1993; Nosofsky, 1984; Simon & Gobet, 1996). The concept of Blending is, however, a recent ACT-R proposition (Lebiere, 1998). This mechanism may explain the process of "intuition" frequently suggested in naturalistic decision-making research (Klein, 1998).

Results from the third experiment series demonstrate better fit to human data when the model waits until the last minute to react to an urgent situation (i.e., low necessity level). IBLT predicts decision makers improve their time of intervention with practice in the task. Therefore, we expect a perception of high urgency for inexperienced decision. With practice in the task, the decision maker's necessity level lowers gradually with learning to wait until the right moment to execute a decision. Empirical research suggests that humans are not very good at selecting the right time of intervention in dynamic tasks (Kerstholt & Raaijmakers, 1997). But more empirical studies are necessary to understand the issues involved in timing of decisions and their relationship to practice in a task. IBLT suggests that there are individual differences in the feeling of urgency and that decision makers learn to adapt their time of intervention according to the similarity of the situations they experience. The analysis of specific individuals indicated best model fit for fast learners compared to slow learners. More research is needed on individual differences as related to DDM. In previous research, we have found that the effects of workload on learning are modulated by the individuals' cognitive capacity (Gonzalez, submitted for publication; Gonzalez, Qudat-Ullah, & Lerch, submitted for publication). Smith, Patalano, and Jonides (1998) suggested that WM determines the use of rules rather than exemplars. Their rationale is that rule application may involve serial processing requiring multiple acts of attention while the use of exemplars may require parallel processing and one-time retrieval from long-term memory. The similarity-rate mechanism proposed in IBLT can be a WM capacity indicator, since it determines the number of SDUs used in the recognition and judgment processes. Also, ACT-R provides forgetting and spreading of activation mechanisms that have been proposed as a source of individual differences and we need to explore further (Lovett, Reder, & Lebiere, 1999).

Results from the fourth series of experiments showed improved performance and learning when accounting for results to upgrade the utility of instances, compared to ignoring feedback. The results also showed that humans don't account for feedback to upgrade instances. Determining the causal relations from outcomes to decisions is a major research topic in DDM (Brehmer, 1990, 1992; Diehl & Sterman, 1995). In the past, it has been shown that humans misperceive feedback due to delays inherent to dynamic situations (Kerstholt & Raaijmakers, 1997). Our results suggest that humans may not be able to attribute an outcome to the right set of decisions that produced it. In DDM, outcomes must be attributed to a sequence of decisions rather than to just one decision. Another possibility is that the knowledge of the results may not affect all the decisions in the same way, as we have assumed in CogIBLT. Some decisions may have a greater effect on the outcomes than others. In general, CogIBLT parameters need to be tested further with sensitivity analyses that could help us understand how unique the learning curves are to the modeling decisions. Also, more empirical research is needed to understand how decision makers account for feedback in dynamic environments.

## 8. Concluding remarks

DDM poses a set of challenges to our understanding of human cognition, capabilities, and limitations. DDM requires recognition, judgement, choice, and re-evaluation of choices in a continuously changing environment. Decision making in DDM is inherently time constrained, highly uncertain due to mainly the exogenous environmental events, highly complex due to multiple and interdependent decisions. In this paper we proposed that, decision making in DDM occurs by the acquisition, retrieval, and refinement of decision-situation-utility instances. We presented a set of learning mechanisms that account for instance-based learning in the context of a decision-making process. The use of a computational model based on ACT-R, makes CogIBLT a precise, predictive tool to experiment with IBLT propositions. We expect IBLT to be a theory provocative enough to instigate psychology and behavioral decision research so as to advance the study of decision making in complex, dynamic situations. Furthermore, we expect CogIBLT to be of interest to researchers in AI and computational modeling of human cognition. Researchers developing human-like decision systems should aim at a better understanding and accurate representations of human learning. Computational models of decision making should reproduce the variability in preferences within single individuals, the adaptability and flexibility that humans exhibit in dynamic situations, and the interrelation between decisions and the influence of results in future choices.

## Notes

1. Elementary Perceiver and Memorizer, originally proposed by Simon and Feigenbaum (1964).
2. Refer to Appendix A and see Anderson and Lebiere (1998), Chapters 3 and 4.
3. We added numbers to the tanks to be able to make references to the picture in this text. These are invisible to the end-user.
4. A detailed description of WPP and performance tests of different heuristics can be found in the Supplemental Materials section through the Cognitive Science on-line Annex at <http://cogsci.psy.utexas.edu/supplements/>.
5. A description of the model's production in an algorithmic form is available in the Cognitive Science on-line Annex at <http://cogsci.psy.utexas.edu/supplements/>.
6. See the on-line Annex for more information on the verbal protocols.
7. Refer to Appendix A for a detailed explanation of Blending.
8. Refer to Appendix A for a detailed explanation of Blending.
9. This update process has not been implemented as a production by itself. It is a LISP process activated in Production 1.
10. At 2:30 the next deadline is 5:00 or 300 min, resulting in time left of  $300 - 138 = 162$ .
11. See the Supplemental Materials section through the Cognitive Science on-line Annex at <http://cogsci.psy.utexas.edu/supplements/>.
12. See Appendix A for the default values of ACT-R parameters used in CogIBLT.
13. Both model and human data averages are over 14 subjects.
14. The average RMSSD for human data only is 9.24.

15. The average RMSSD for human data only is 9.24.
16. We never asked our participants to follow any rule. And even if we had instructed them to follow a rule, there is no sure way of knowing that participants are in fact following a rule.
17. Further explorations of the time rule fit and similarity for these two participants can be found in the Supplemental Materials section through the Cognitive Science on-line Annex at <http://cogsci.psy.utexas.edu/supplements/>.
18. Further explorations of the time rule fit and similarity for these two participants can be found in the Supplemental Materials section through the Cognitive Science on-line Annex at <http://cogsci.psy.utexas.edu/supplements/>.
19. If alternatives are retrieved in order of cue importance.

#### Acknowledgments

This research was supported by the Air Force Office of Scientific Research (F49620-97-1-0368), the Advanced Decision Architectures Collaborative Technology Alliance sponsored by the U.S. Army Research Laboratory (DADA19-01-2-0009), and the Office of Naval research (N00014-96-1-0322).

#### Appendix A. ACT-R mechanisms and parameters

Activation in ACT-R is calculated according to Bayesian methods as:

$$A_i = B_i + \sum_j W_j S_{ji}$$

where  $B_i$  is the base-level activation, an estimate of how likely a chunk will match to a production. The  $B_i$  value depends on the frequency and latest chunk usage.  $W_j$  is the attention given to source  $j$  and  $S_{ji}$  is the strength of association from a cue  $j$  to a chunk  $i$ .

PM allows the retrieval of a chunk even when it only partially matches a production condition. The basic idea is to define a similarity metric over the range of possible values. This can be done whether those values are continuous or discrete. In the case of discrete values, this can be done by simple enumeration of similarities between pairs of values. For continuous values (e.g., real numbers), one can define a similarity function over that domain. One can then combine the activation of a chunk with its degree of match to the production condition to yield a composite measure of fit called the match score  $M_{ip}$  of chunk  $i$  to production condition  $p$ :

$$M_{ip} = A_i - MP \sum_{u,d} (1 - \text{Sim}(u, d)) \quad \text{Partial Matching Equation}$$

The degree of mismatch for a production condition is defined as  $1 - \text{sim}$ , i.e., the maximum possible similarity, minus the similarity  $\text{Sim}(u, d)$  between desired value  $d$  specified in the production condition and the actual value  $u$  present in the chunk. The match score is computed by subtracting from the chunk's activation the sum of all the conditions specified by the production

multiplied by the scaling parameter MP. That parameter can be used to trade off activation and degree of match. If it is equal to 0, then the degree of match to production condition is irrelevant and one simply retrieves the most active chunk that comes to mind. If the parameter gets very large, then degree of match becomes paramount and this becomes the standard symbolic exact matching. For intermediate values, the chunk with the highest match score will be retrieved from memory if its score is above the activation threshold  $\tau$ . Otherwise, the chunk retrieval fails.

Blending is a generalization of PM that allows the retrieval of an aggregate result of a set of memory chunks rather than only one chunk (Lebiere, 1998). Instead of retrieving a single chunk from declarative memory, the aggregate result of a set of memory chunks is retrieved, where aggregate is defined in terms of the match scores and similarities of the PM mechanism. The idea is that each chunk proposes a value to be retrieved with a strength reflecting the match score of the chunk. The aggregate answer  $V$  is defined as the value that minimizes the sum of the squared dissimilarities with the answer  $V_i$  proposed by each chunk, weighted by the chunk's probability of retrieval  $P_i$ :

$$V = \text{Min} \sum_i P_i (-\text{Sim}(V, V_i))^2 \quad \text{Blending Equation}$$

As in the Partial Matching Equation,  $\text{Sim}(V, V_i)$  is the similarity between value compromise value  $V$  and actual value  $V_i$  returned by chunk  $i$ .  $P_i$  is the probability of retrieving chunk  $i$  as a function of its match score  $M_i$  and the match scores  $M_j$  of all other chunks  $j$ , given by the Boltzmann (aka. softmax) equation (Anderson & Lebiere, 1998):

$$P_i = \frac{e^{M_i/\tau}}{\sum_j e^{M_j/\tau}} \quad \text{Retrieval Probability Equation}$$

where  $\tau$  is a measure of the activation noise. The lower the noise, the more the system will behave deterministically in retrieving the chunk with the highest match score. The higher the noise, the more the system will randomly retrieve any chunk that partially matches.

For instance, for the values in Fig. 5, the similarity between the current time deadline (128) and water amount (3.19) and the corresponding values in the three final chunks is combined with their activations to compute their match score according to the Partial Matching Equation. The match scores are then combined to yield each chunk's retrieval probability using the Retrieval Probability Equation. The utility values proposed by those chunks (306, 309, and 293) are then combined using the Blending Equation to yield a composite answer, 297.83. Blending applies equally well to continuous and to discrete domains. At the discrete end of the spectrum, if no similarities are defined between values, then each chunk proposes its own answer and the strongest wins. This is equivalent to the current retrieval mechanism, with the exception that separate chunks proposing the same answer can now pool their strengths instead of competing separately, providing a dynamic generalization of the ACT-R chunk merging mechanism in which identical chunks are merged together and their activation strengths are combined. At the continuous end of the spectrum, the answer is the average of the values proposed by each chunk, weighted by their probabilities of retrieval. Intermediate points along the spectrum include integers, coarse scales, and domains with some similarities but no regular scale. Blending can

be viewed as a generalization of well-known AI techniques. Neural networks have a similar ability to learn in their connection weights a number of training patterns and produce an output that reflect the constraints of the entire training set rather than any specific pattern. The Bayes Optimal Classifier produces the most likely outcome weighted over all hypotheses (ACT-R chunks), rather than simply the most likely hypothesis (most active chunk). Linear weighted regression is an instance-based machine learning algorithm that produces the answer that minimizes the squared error between a fitted function and a set of data points, with each data point being weighted by its distance to the query point. The Blending mechanism combines attributes of all these techniques.

The following list shows the ACT-R parameter values used as default values in CogBLT throughout the experiments.

Enable rational analysis (era): 1  
*Gf*(*g*): 20.0  
 Expected gain *S* (egs): nil  
 Enable randomness (er): nil  
 Utility threshold (ur): 0  
 Goal activation (ga): 0  
 Base-level constant (ble): 2.5  
 Activation noise *S* (ans): 0.25  
 Permanent activation *S* (pas): nil  
 Latency factor (lf): 1.0  
 Latency exponent (le): 1.0  
 Default action time (dat): 0.05  
 Partial Matching (PM): 1  
 Mismatch penalty (mp): 1.5  
 Retrieval threshold (rt): 0  
 Optimized learning (ol): 1  
 Base-level learning (bl): nil  
 Associative learning (al): nil  
 Strength learning (sl): nil  
 Parameters learning (pl): nil

## Appendix B. Similarity calculation

Consider for example the decisions made by participant x in Tank 0. The similarity of the SDUs in Trial 1 will be 0. This participant made three decisions in Trial 2 on Tank 0. Four decisions on Trial 3, etc. The similarity of the first SDU in Trial 2 results in a value of 0.5 as follows:  $1 - (0.99 \times ((172 + 15 \times 2) - (390 + 27 \times 2)) + 0.01 \times 0)$ . The table shows how the similarity value is calculated for several SDUs.

User id	Tank	Trial	Order	Time	Water	Similarity	Similarity calculation
x	0	1	1	390	27	0	
x	0	2	1	172	15	0.50	$1 - (0.99 \times \text{ABS}((172 + 15 \times 2) - (390 + 27 \times 2)) + 0.01 \times 0)$
x	0	2	2	226	2	0	
x	0	2	3	285	8	0	
x	0	3	1	147	10	0.92	$1 - (0.99 \times \text{ABS}((147 + 10 \times 2) - (172 + 12) + 480 + 0.01 \times 0.5))$
x	0	3	2	222	2	0.99	$1 - (0.99 \times \text{ABS}((222 + 2 \times 2) - (226 + 2 \times 2) + 480 + 0.01 \times 0))$
x	0	3	3	248	2	0.90	$1 - (0.99 \times \text{ABS}((248 + 2 \times 2) - (285 + 8 \times 2) + 480 + 0.01 \times 0))$
x	0	3	4	269	2	0	

## References

- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum (Lawrence).
- Anderson, J. R., Greene, J. G., Kline, P. K., & Neves, D. M. (1981). Acquisition of problem solving skill. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale, NJ: Erlbaum (Lawrence).
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Hillsdale, NJ: Erlbaum (Lawrence).
- Berry, D., & Broadbent, D. A. (1984). On the relationship between task performance and associated verbalized knowledge. *The Quarterly Journal of Experimental Psychology*, 36A, 209-231.
- Brehmer, B. (1990). Strategies in real-time, dynamic decision making. In H. R. M. (Ed.), *Issues in decision making* (pp. 262-279). Chicago: University of Chicago Press.
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Psychologica*, 87, 211-231.
- Carr, T. H., McCauley, C., Sperber, R. D., & Parmelee, C. M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and automaticity of information processing. *Journal of Experimental Psychology*, 111, 757-777.
- Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.
- Cohen, A. L., & Nisolsky, R. M. (2000). An exemplar-retrieval model of speeded same-different judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 26(5), 1549-1569.
- de Groot, A. D. (1978). *Thought and choice in chess* (2nd ed.). The Hague: Mouton.
- Diehl, E., & Sternan, J. D. (1995). Effects of feedback complexity on dynamic decision making. *Organizational Behavior and Human Decision Processes*, 62(2), 198-215.
- Domingos, P. (1996). Unifying instance-based and rule-based induction. *Machine Learning*, 24, 141-168.
- Dreyfus, H. L. (1997). Intuitive, deliberative, and calculative models of expert performance. In C. E. Zsombok & G. Klein (Eds.), *Neuralistic Decision Making* (pp. 17-28). Mahwah, NJ: Erlbaum (Lawrence).
- Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: The Free Press.
- Edwards, W. (1962). Dynamic decision theory and probabilistic information processing. *Human Factors*, 4, 59-73.
- Forbus, K. D., Gentner, D., & Law, K. (1994). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19, 141-205.
- Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gibson, L., & Schneider, D. (1995). Case-Based Decision Theory. *The Quarterly Journal of Economics*, 110, 605-639.
- Gibson, L., & Schneider, D. (1997). Act similarity in Case-Based Decision Theory. *Economic Theory*, 9, 47-61.



- Gilboa, I., & Schmeidler, D. (2000). Case-based knowledge and induction. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30(2), 85-95.
- Gonzalez, C. *Decision making in a dynamic, real-time environment: Learning under high cognitive workload*. Manuscript submitted to Human Factors for publication.
- Gonzalez, C., Qudus-Ullah, H., & Leich, F. J. *Decision making in a dynamic, real-time environment: Learning under time pressure*. Manuscript submitted to Human Factors for publication.
- Haider, H., & Freinsch, P. A. (1996). The role of information reduction in skill acquisition. *Cognitive Psychology*, 30, 304-337.
- Hartmann, D. L. (1984). MINVERA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, 16, 96-101.
- Hartmann, D. L. (1986). "Schema Abstraction" in a multiple-trace memory model. *Psychological Review*, 93(4), 411-438.
- Kerstholt, J. H., & Raaijmakers, J. G. W. (1997). Decision making in dynamic task environments. In R. Ranjard, R. W. Crozier, & O. Svenson (Eds.), *Decision making: Cognitive models and explanations* (pp. 205-217). New York, NY: Routledge.
- Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.
- Klein, G., Orasanu, J., Calderwood, R., & Zsombok, C. E. (Eds.). (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex Publishing Corporation.
- Lassaline, M. L., & Logan, G. (1993). Memory-based automaticity in the discrimination of visual numerosity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 561-581.
- Lebiere, C. (1998). *The dynamics of cognition: An ACT-R model of cognitive arithmetic* (CNU-CS-98-186. Available at: <http://reports-archive.adm.cs.cmu.edu/>). Pittsburgh, PA: Carnegie Mellon University.
- Lebiere, C. (1999). The dynamics of cognitive arithmetic. *Kognitionswissenschaft Journal of the German Cognitive Science Society*, 8, 5-19 (1). Special issue on cognitive modeling and cognitive architectures.
- Lebiere, C., & Bobell, D. (2002). *Multi-tasking and cognitive workload in an ACT-R model of a simplified air traffic control task*. Paper presented at the 10th Conference on Computer Generated Forces and Behavioral Representation, Norfolk, VA.
- Lebiere, C., Wallach, D., & Niels, T. (1998). *Implicit and explicit learning in ACT-R*. Paper presented at the Second European Conference on Cognitive Modeling, Groningen, The Netherlands.
- Leinard, J. W. (2000). *Similarity judgments, violations of stationarity, and reflection effects in intertemporal choice* (RC 21893 (98504)). Yorktown Heights: IBM Research Division, Thomas J. Watson Research Center.
- Leitch, F. J., Balbou, D. B., & Harter, D. E. (1997). Using simulation-based experiments for software requirements engineering. *Annals of Software Engineering*, 3, 345-366.
- Logan, G. (1988). Toward an Instance Theory of Automation. *Psychological Review*, 95(4), 492-527.
- Logan, G. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, 22, 1-35.
- Logan, G. (1992). Shapes of reaction time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 883-914.
- Logan, G., & Klapp, S. T. (1991). Automating alphabet arithmetic: Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 179-195.
- Lovett, M. C., Reder, L. M., & Lebiere, C. (1999). Modeling working memory in a unified architecture: An ACT-R perspective. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control*. New York, NY: Cambridge University Press.
- Marcel, A. T. (1983). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology*, 15, 238-300.
- March, J. G. (1994). *A primer on decision making: How decisions happen*. New York: The Free Press.
- Markman, A. B., & Medin, D. L. (1995). Similarity and alignment in choice. *Organizational Behavior and Human Decision Processes*, 63(2), 117-130.
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, 100(2), 254-278.
- Medin, D. L., Goldstone, R. L., & Markman, A. B. (1995). Comparison and choice: Relations between similarity processing and decision processing. *Psychonomic Bulletin and Review*, 2(1), 1-19.

- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, 85(3), 207-238.
- Medin, D. L., Watermark, W. D., & Michalski, R. S. (1987). Constraints and preferences in inductive learning: An experimental study of human and machine performance. *Cognitive Science*, 11, 299-339.
- Merrill, E., Sun, R., & Peterson, T. (2001). From implicit skills to explicit knowledge: A bottom-up model of skill learning. *Cognitive Science*, 25, 203-244.
- Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(1), 104-114.
- Nosofsky, R. M., & Palmer, T. J. (1997). An Exemplar-Based Random Walk model of speeded classification. *Psychological Review*, 104(2), 266-300.
- Palmer, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(2), 324-354.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York, NY: Cambridge University Press.
- Pew, R. W., & Mavor, A. S. (1998). *Modeling human and organizational behavior*. Washington: National City Press.
- Ram, A. (1993). Indexing, elaboration and refinement: Incremental learning of explanatory cases. *Machine Learning*, 10, 201-248.
- Reisbeck, C. K., & Schank, R. C. (1989). *Inside case based reasoning*. Hillsdale, NJ: Erlbaum (Lawrence).
- Rubinstein, A. (1988). Similarity and decision-making under risk. *Journal of Economic Theory*, 46, 145-153.
- Samet, S., Anderson, J. R., Lebiere, C., & Lovett, M. C. (2000). *Achieving efficient and cognitively plausible learning in backgammon*. Paper presented at the 17th International Conference on Machine Learning (ICML-2000), Stanford, CA.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, 84, 1-66.
- Schunn, C., & Wallach, D. (2002). *In defense of goodness-of-fit in comparison of models to data*. Unpublished manuscript, Pittsburgh, PA.
- Simon, H. A. (1957). *Models of man: Social and rational*. New York: Wiley.
- Simon, H. A., & Amaze, Y. (1979). The theory of learning by doing. *Models of thought* (Vol. II, pp. 116-133). New Haven, London: Yale University Press.
- Simon, H. A., & Feigenbaum, E. A. (1964). An information-processing theory of some effects of similarity, familiarization, and meaningfulness in verbal learning. *Journal of Learning & Verbal Behavior*, 3(5), 385-396.
- Simon, H. A., & Gabel, F. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Science*, 31, 1-40.
- Simon, H. A., & Langley, P. (1981). The central role of learning in cognition. In H. A. Simon (Ed.), *Models of thought* (Vol. II, pp. 167-184). New Haven, London: Yale University Press.
- Simon, H. A., & March, J. G. (1958). *Organizations*. New York: Wiley.
- Simon, H. A., & Zhu, X. (1988). Learning mathematics from examples and by doing. In *Models of thought* (Vol. II, pp. 167-184). New Haven, London: Yale University Press.
- Smith, E. E., Palano, A. L., & Jones, J. (1998). Alternative strategies of categorization. *Cognition*, 65, 167-196.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Tversky, A. (1977). Features of similarity. *Psychology Review*, 84, 337-352.
- Watson, I. (1995). An introduction to Case-Based Reasoning. In I. Watson (Ed.), *Progress in Case-Based Reasoning* (p. 209). Salford, UK: Springer.
- Zsombok, C. E., & Klein, G. (Eds.). (1997). *Naturalistic Decision Making*. Mahwah, NJ: Erlbaum (Lawrence).