

# Learning to Make Decisions in Dynamic Environments: Effects of Time Constraints and Cognitive Abilities

Cleotilde Gonzalez, Carnegie Mellon University, Pittsburgh, Pennsylvania

This study investigated the effects of time constraints and cognitive abilities on dynamic decision making (DDM). The learning and performance of individuals trained in a DDM task with time constraints were compared with those who were trained without time constraints. Although all participants received the same total amount of time to perform the task, individuals under more stringent time constraints were given 3 times more practice trials on the first 2 days of the study than were people under less stringent time constraints. Despite the additional practice runs, participants under high time constraints performed worse than did participants under low time constraints on the 3rd day of the study. A subsequent analysis of cognitive abilities and decision heuristics revealed that individuals' actions corresponded with simple heuristic predictions more closely with minimal practice than with extended practice, under high rather than low time constraints, and in individuals with low rather than high cognitive abilities. These findings suggest that mere repetition of a task with less time within a trial is counterproductive to learning and that learning depends on cognitive abilities. Potential applications of this research include the design of training procedures for dynamic tasks.

## INTRODUCTION

Dynamic decision making (DDM) comprises a series of multiple, interdependent decisions that are made in real time in a continuously changing autonomous environment (Brehmer, 1990; Edwards, 1962). For example, air traffic control (ATC) requires controllers to make multiple decisions regarding how to allocate space to best accommodate the arrivals and departures of airplanes. That the assignment of a landing lane to an incoming airplane precludes the use of that lane by other airplanes, reflects the interdependency of decisions that characterizes DDM tasks. Furthermore, environmental parameters such as arrivals, departures, and weather are exogenous during ATC (i.e., they are beyond the influence of the controller). Finally, incoming airplanes need to be assigned to a landing lane at the correct moment in real time. Thus ATC provides a prime real-world example of DDM.

In actuality, most real-life decisions involve

DDM, although they vary in the time allowed for decision making and in the number of chances that one receives to practice them. For example, in an effort to become tenured, professors make multiple and interrelated decisions regarding how to best allocate their time to tasks such as manuscript preparation, grant writing, and teaching. Throughout this process, many exogenous factors such as teaching load and reviewers affect the professors' decision-making processes, and the time at which decisions are made is critical for the accomplishment of the professors' motivating goal.

*Time constraints* have been defined as the difference between the amount of available time and the amount of time required to resolve a decision task (Benson & Beach, 1996; Rastegary & Landy, 1993). According to this definition, time constraints are relative to the pace of change in the decision environment. For example, professors in pursuit of tenure have much more time to make decisions than do air traffic controllers

trying to help airplanes land safely. When placed under time constraints, decision makers encounter multiple decisions per unit time. Time constraints and the number of task executions are factors that influence individuals' ability to acquire control of a dynamic decision system (Kerstholt & Raaijmakers, 1997). Intuitively, one might expect that the more times individuals practice a task the better their task performance will be, but the interaction of time constraints and practice and the influence of these factors on performance have not been studied. This study addressed this research question in a DDM environment by examining both the relationship between practice trials and time constraints and the possible cognitive strategies utilized by individuals with different cognitive abilities.

#### **INFLUENCE OF TIME AND COGNITIVE ABILITIES ON DDM**

Research has shown that time constraints have a negative effect on the ability of individuals to make decisions effectively (Maule & Edland, 1997; Svenson & Maule, 1993). Most of this research, however, has emphasized traditional, static decision-making tasks and has focused almost exclusively on the effects of time pressure on performance. Researchers also have primarily investigated one-time decisions rather than the series of decisions that characterize DDM environments (Kerstholt & Raaijmakers, 1997).

Although most research conducted to date has evaluated the factor of time constraints on individuals performing static decision-making tasks, there is no reason to expect this effect to be anything other than detrimental to the ability of individuals to perform dynamic tasks as well. Because many DDM situations are extremely complex, the amount of information that an individual must process before making a decision can be very taxing on the decision maker if time is limited (i.e., if the decision maker is working under a deadline or within a higher-paced environment). As in static situations, individuals asked to perform dynamic tasks may exercise "satisficing" – thus limiting the amount of information they must process before making a decision – or may simply reduce their efforts at processing to perform the task as quickly as possible. In any case, time constraints would be expected to induce lower performance.

It is more difficult to predict what sort of effect time constraints might have on learning. One might reasonably hypothesize that the ability to deal with time constraints effectively improves with practice. If this is the case, the completion of enough practice trials should help individuals to achieve control over a system and therefore should improve overall task performance. Hence, with sufficient practice, individuals under severe time constraints should reach performance levels similar to those of individuals under low time constraints. In summary, this study was designed to test the hypothesis that high time constraints attenuate dynamic task performance and that practice can moderate this negative effect.

Human cognitive capacity is another variable that is frequently mentioned in the literature but rarely investigated in DDM studies. Because human cognition is limited, it should modulate the effects of time pressure on learning in DDM situations. In complex and dynamic systems, the need to process larger amounts of information increases as the amount of available time decreases. Because dynamic situations change over time, decision makers must process new situations continuously and be able to retain information while concurrently processing incoming variables. Researchers commonly define *fluid intelligence* (Gf) as the ability to solve novel problems and adapt to new situations (Engle, Tuholski, Laughlin, & Conway, 1999). Working memory and Gf are related through what is commonly referred to as *controlled attention*, or the number of elements on which an individual can focus his or her attention at a particular time. The study presented here examined the possible effects of human cognitive capacity and its relationship to time constraints. Because human capacity is limited, it was hypothesized that the detrimental effects of time constraints would be substantial and would correlate with a measure of cognitive capacity.

This study also investigated the effects of time constraints on the strategies that individuals use to deal with a dynamic system. In static situations with time constraints, decision makers speed up their information processing and reduce the amount of time they spend searching for predecision information (Edland & Svenson, 1993). Kerstholt (1994) reported that during DDM there is an acceleration in information

processing as time becomes more limited. She also found that dynamic decision makers employ judgment-oriented strategies (i.e., searching for information before choosing to perform an action) rather than action-oriented strategies (i.e., performing an action and justifying it afterward), but her studies provided no insight into how these decision strategies change when individuals are placed under time constraints (Kerst-holt, 1996).

Research on decision making under conditions of uncertainty suggests that decision makers react to time constraints by applying simple heuristics (experience-based rules) more quickly and avoiding strategies that require too much information processing (Gigerenzer & Todd, 1999). Thus one might reasonably expect to observe a greater use of heuristics by dynamic decision makers under high time constraints than by those under low time constraints. However, it is unclear how the use of heuristics changes over time with task practice. Some researchers claim that the use of heuristic methods is one of the most important end products of experience (Gigerenzer & Todd, 1999), whereas others suggest that decision makers use experience-based strategies and, in so doing, ignore general algorithms in order to choose specific solutions recognized through experience ("context-based solutions"; Dreyfus, 1997; Logan, 1988).

This study tested the hypothesis that decision makers use simple heuristics during early encounters with a DDM task but move away from these general solutions to context-based solutions as they learn to perform the task more adeptly. The learning experiment was designed to evaluate whether heuristics are used by dynamic decision makers under low or high time constraints and, further, whether the use of heuristics varies

with practice in the task. The putative interaction between other task constraints and cognitive abilities is difficult to delineate and is the focus of other research (Gonzalez, in press). However, one might reasonably expect that cognitive limitations of individuals under high time constraints might have a negative effect on performance improvement.

## METHOD

### Experimental Design

The experimental design is shown in Table 1. All participants ran a DDM simulation (described later) on 3 consecutive days. The results of a pilot study were used to define two levels of time constraints. Under the fast condition, each simulation trial lasted 8 min. Participants under this condition completed 18 trials over the 3-day period (6 trials/day). Under the slow condition, each simulation trial on the first 2 days lasted 24 min (2 trials/day), whereas each trial on the last day lasted 8 min (6 trials). For all participants, the first 2 days were practice days and the last day was the test day.

This design enabled an assessment of the extent to which participants were able to transfer the practice knowledge they gained under low time constraints to the testing environment under high time constraints. All participants had the same total time on task (144 min) and the same time on task per day (48 min). However, participants in the fast-condition group performed the task 18 times (all in fast mode), whereas participants in the slow-condition group completed a total of 10 trials (4 trials at the slow pace and 6 trials in fast mode). Thus the design also allowed testing of the hypothesis that people who practice a DDM task under low time constraints will perform better under high time

TABLE 1: Experimental Design

	Day 1	Day 2	Day 3	No. of Trials	Time on Task (min)
Fast	6 trials of 8 min each	6 trials of 8 min each	6 trials of 8 min each	18	144
Slow	2 trials of 24 min each	2 trials of 24 min each	6 trials of 8 min each	10	144
Time on task	48 min	48 min	48 min		

constraints than will people who practice consistently under high time constraints.

### Participants and Cognitive Abilities

Thirty-three graduate and undergraduate college students recruited from local universities were randomly assigned to either the fast- or the slow-condition group. They were paid \$50 at the end of the 3-consecutive-day evaluation period.

Cognitive abilities were measured by using the Raven Standard Progressive Matrices Test (Raven, Court, & Raven, 1977). This test was selected because it is a well-known, validated measure of Gf, does not depend on participants' verbal abilities, and is free of cultural bias. Psychological research suggests that the Raven test is a useful predictor of the ability to perform dynamic tasks because the test measures the ability of people to manage a large set of goals in their working memory, to solve novel

problems, and to adapt to new situations (Carpenter, Just, & Shell, 1990; Engle et al., 1999). The Raven test consists of a set of visual analogy problems. Each problem presents, at the top of the page, a pattern in which a figure is missing. The test asks the participants to select the option that best completes the background. The test comprises five sets of 12 questions each for a total of 60 questions, which are arranged according to degree of difficulty (with the more difficult questions presented at the end). The participant's score is the total number of correct answers (possible range: 0–60). This test takes approximately 40 min to complete.

### Dynamic Decision-Making Task

The DDM simulation used for this study, the Water Purification Plant™ (WPP; Figure 1), is an isomorph of a real-world scheduling task performed in an organization with large-scale logistical operations (Lerch, Ballou, & Harter, 1997).

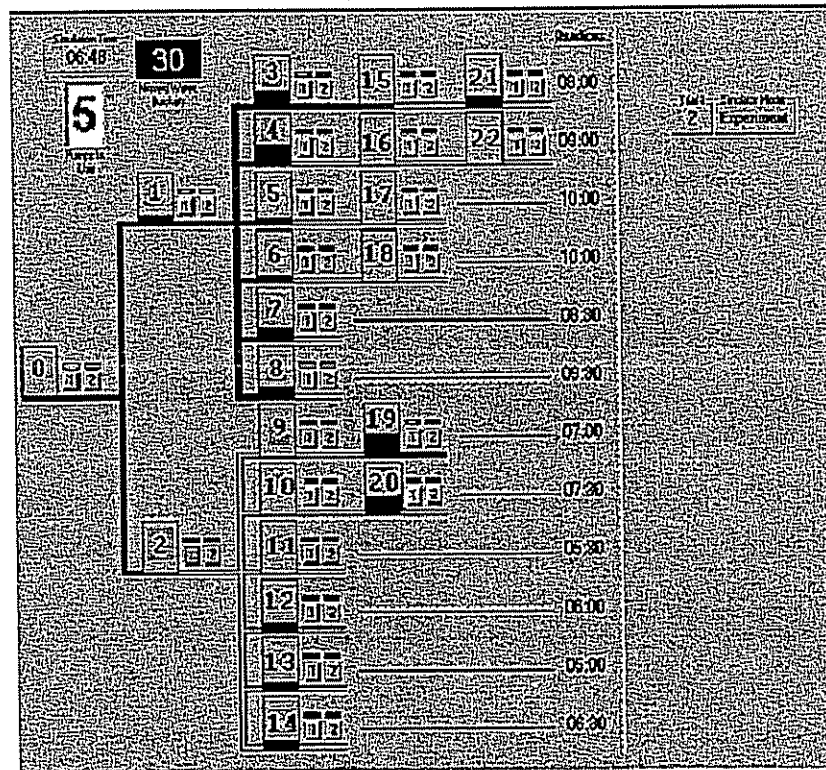


Figure 1. Layout of the WPP simulation. The simulation time is 6:48 p.m. The operator has missed 30 gallons of water, and five pumps (the maximum number allowed) currently are activated. Water enters from outside the system and moves continuously through the activated pumps from left to right toward the deadlines. The operator decides when to activate and deactivate pumps as the simulation time runs continuously.

The WPP program simulates a water distribution system with 23 tanks arranged in a tree structure and connected with pipes. The goal in this task is to distribute all the water in the system before the various deadlines expire (see right side of Figure 1). Users distribute the water by activating and deactivating pumps located next to each tank. The simulation runs from 2:00 to 10:00 (8 simulation hr), labeled as one trial. Consecutive tanks, such as Tank 3, Tank 15, and Tank 21, have the same deadline, 8:00 p.m. The root tanks take the earliest deadline – that is, the deadline for Tanks 0 and 2 is 5:00 p.m., and the deadline for Tank 1 is 8:00 p.m. Each tank has a fixed position in the tree structure (called the *chain value*) that indicates the number of tanks through which the water needs to travel; Tank 22 has a chain of one (only one tank needs to be emptied before the deadline), Tank 4 has a chain of three, Tank 0 has a chain of five, Tank 1 has a chain of four, and so forth. The higher the chain, the longer it takes to deliver the water. The system processes 1 gallon (3.785 L) of water every 2 simulation min for each pump. Thus, using two pumps makes the processing rate twice as fast (1 gallon of water/simulation min).

The WPP simulation fulfills the primary criteria of a DDM task (Brehmer & Dörner, 1993). The environment is opaque, which leads to user uncertainty as to the values of some key variables. (For example, water appears in the system according to a scenario defined by the experimenter and unknown to the user.) The environment changes autonomously and in response to the user's decisions. Because a maximum of five pumps can be activated at any one time, the decision maker's actions are interrelated. For example, a user's decision to activate a pump in one chain may preclude adjusting the water flow in another. Furthermore, there is a delay between the time at which a pump is deactivated and the time at which it can be reactivated. This idle time decreases the amount of time available for the user to meet the deadlines. More details on this task are available elsewhere (Gonzalez, Lerch, & Lebiere, 2003).

### Performance Measure

WPP performance is measured as the gallons of water that the user fails to distribute on time.

The system has a total capacity of 1080 gallons (4088 L) of water. Optimal performance earns a score of zero (no water missed). The number of missed gallons was converted into a positive percentage measure using these values. Thus, large percentages (i.e., lower numbers of missed gallons) indicate better performance. There are many possible decision sequences by which a user can activate and deactivate the pumps to achieve optimal performance. To serve as a reasonable lower limit for performance, a program was created to run the simulation while randomly activating the pumps and maintaining no idle time (i.e., never having fewer than five pumps activated). Thirty trials conducted by this random simulation program produced a mean of 182.9 missed gallons (81.9%; 692.3 missed L) with a standard deviation of 28.4 (2.4%). Therefore, for the purposes of this study a score of 0 to 200 (0%–80%) constituted a reasonable first-time performance.

### Process Variables

According to verbal protocols conducted previously on the WPP (Gonzalez, 2003), individuals use the simulation time, the deadlines, the amount of water in the tanks, and the chain value to determine which tank to activate. Three decision heuristics were identified accordingly: time (T), time-volume (TV), and time-volume-chain (TVC).

The T heuristic suggests activating a pump on a tank with the closest deadline to the current simulation time. For example, if the simulation time is 3:00, the T heuristic suggests that the user activate the pumps for Tank 13 (see Figure 1), for which the deadline is 5:00 (only 120 min are left for the deadline), rather than the pumps for Tank 11 (for example), for which the deadline is 5:30 (150 min are left).

The TV heuristic finds the most urgent pump to operate by assessing the amount of water in the tanks in addition to the time left prior to the deadlines. A tank can hold up to 100 gallons (378.5 L) of water, and since the system processes water at a constant speed of 1 gallon every 2 simulation min, it is possible to estimate how long it takes to empty a tank and calculate the urgency to act on a tank. According to the TV heuristic, tanks that contain a large amount of water but are under no immediate deadline

would be activated before tanks containing little water that are under an immediate deadline. For example, if the simulation time is 5:00, the amount of water in Tank 19 is 60 gallons (227 L) and the amount of water in Tank 12 is 10 gallons (37.85 L; Figure 1), the TV heuristic would recommend activating Tank 19, despite the delayed deadline, since the time left to process the water in this tank,  $120 - (60 \times 2) = 0$  min, is less than the time left to process the water in Tank 12,  $60 - (10 \times 2) = 40$  min.

Finally, the TVC heuristic is similar to the TV heuristic except that it also considers the chain position of the tank. For example, at 5:00, if Tank 1 and Tank 21 in Figure 1 held the same amount of water (assume 20 gallons, or 76 L, of water), the TVC heuristic would recommend activating Tank 1 rather than Tank 21 because the water in Tank 1 would take longer to pump through the system because of its chain location,  $[180 \text{ min} - (20 \times 2)] \times 4 = 560$  min, than Tank 21,  $[180 - (20 \times 2)] \times 1 = 140$  min, and thus the user would not meet the 8:00 deadline.

Decision-making research often compares decision behavior with decisions based on hypothetical heuristics to draw conclusions regarding the decision-making process (Hogarth & Spyros, 1981). This research created the *fit to the heuristic* as a process measure. This measure is similar to that used by other researchers (Payne, Bettman, & Johnson, 1993), except that the fit to the heuristic in the WPP depends on values that change dynamically over the course of the simulation (e.g., water and time). The fit value was calculated by comparing the urgency of each of the participant's decisions according to a heuristic and the urgency of the decision prescribed by the same heuristic. A program processed the decisions made by participants after the data collection. This program compared decisions with each of the three heuristics described earlier. The heuristics program calculated a fit value for each decision using the following formula:

$$fit = \frac{[(actual\ decision - worst\ decision)]}{(best\ decision - worst\ decision)}. \quad (1)$$

In this formula, the actual decision represents the calculation of the heuristic for the decision made by the participant, and the worst and best

decisions are obtained by selecting the lowest and highest values among the calculations of the heuristic fits for all the tanks at the moment of the user's decision. The fit is a value that ranges from 0 (no fit) to 1 (perfect fit to the heuristic, or the participant did what the heuristic prescribes). A continuous value from 0 to 1 was used rather than a binary fit value (0 or 1) to provide a more continuous and, possibly, more sensitive metric. Although a binary fit value might be more straightforward, a continuous value measures partial fits to a heuristic.

As an example, assume the participant had just activated the pump in Tank 3, Figure 1. This decision (the actual decision in Equation 1) and the status of all the tanks (amount of water and status of the pumps) were recorded by the WPP at the moment the decision was made. The heuristics program selects the tank to activate according to the heuristic (the best decision in Equation 1). In Figure 1, using the T heuristic, the best decision is Tank 19 because the current time is 6:48 and 7:00 is the closest deadline (Tank 9, because it does not have water to pump, is not considered as an alternative.). Also using the T heuristic as an example, the worse tanks to activate in this situation are those with a 10:00 deadline (worst decision in Equation 1, Tanks 5 and 6). The heuristics program thus would calculate the fit to the T heuristic for this particular decision using the time of the deadlines as follows:

$$fit = [(8 - 10)/(7 - 10)] = .66. \quad (2)$$

That is, making such decision (activating the pump in Tank 3 in the conditions presented in Figure 1) will produce a fit of 66% to the T heuristic. (The TV heuristic converts the volume of water to time units, using the pumping rate of 1 gallon every 2 min. The TVC heuristic in addition multiplies the result of TV by the value of the chain, to reflect the position of the tank in the tree structure.)

Programs created to calculate task performance when each of these heuristics is applied consistently throughout the simulation trial produced the following performance scores: 122 gallons (462 L; 87.9%) for the T heuristic, 99 gallons (375 L; 90.2%) for the TV heuristic, and 75 gallons (284 L; 92.6%) for the TVC heuristic.

Consistent application of any of these heuristics produced better performance scores than did the random strategy (182.9 gallons [692 L], 81.9%). Consistently following the TVC heuristic, which accounts for most of the cues relevant to the task, led to the best performance, whereas consistent application of the simplest heuristic (T) led to the poorest performance among the three heuristic strategies.

### Procedure

The Raven test was administered before the participants began the WPP simulation. After completing the Raven test, participants received instructions on the goals of the WPP simulation and on how to use the simulation to perform the requisite tasks. Instructions were provided through the use of a standard script, and the participants then were permitted to run the simulation once in a training mode (at the very slow speed of 30 min per trial, and only up to the first deadline). During the instructions, participants were informed about the deadlines, simulation time, and water travel paths but were not instructed regarding the use of heuristics. Participants were told that water could enter any of the tanks at any time, but they were not given information about the amount of water to process or the time of water arrival. Participants were instructed to do their best to process all the water that appeared in the system but were taught no strategies. After receiving the instructions, the participants were randomly assigned to either the fast- or the slow-condition group. They ran the simulation for 48 min on each of 3 consecutive days. All experimental trials had the same goal and involved the same scenario.

### RESULTS

Repeated-measures analyses of variance (ANOVAs) using the Raven score as a covariate were used for all statistical analyses. The average Raven score was 53.6, with a standard deviation of 4.2. The Raven score was used as a covariate in the statistical analysis. The average performance per day for each participant was calculated, resulting in three repeated measures per participant (one for each experimental day). The average performance of each fast-

condition participant was calculated by averaging the performance in six trials per day, whereas the average performance of each slow-condition participant was calculated by averaging the performance in two trials on each of the first 2 days and six trials on the 3rd day.

There was a significant main effect of Raven,  $F(1, 29) = 13.20, p = .001$ , and two significant interactions – Day  $\times$  Time Constraints,  $F(2, 58) = 6.67, p = .002$ , and Day  $\times$  Time Constraints  $\times$  Raven,  $F(1, 58) = 6.67, p = .002$  – but no main effect of time constraints. Figure 2 depicts average performance per day by participants under the two different time constraint conditions. On the 1st day, performance was very similar for the fast- and slow-condition groups. Average performance on the 1st day was close to the zero-intelligence scheduler (182.9 gallons [692 L], 81.9%). However, slow-condition participants improved their performance more on the 2nd and 3rd days than did fast-condition participants. As stated earlier, slow-condition participants completed only 4 trials on the 2 practice days (2 trials per day), whereas fast-condition participants practiced the simulation 12 times (6 times per day).

Thus, despite the fact that participants in the slow-condition group had practiced the task fewer times than had participants in the fast-condition group, they outperformed the fast-condition participants by the 2nd day of the study. Furthermore, participants in the slow-condition group continued to improve their

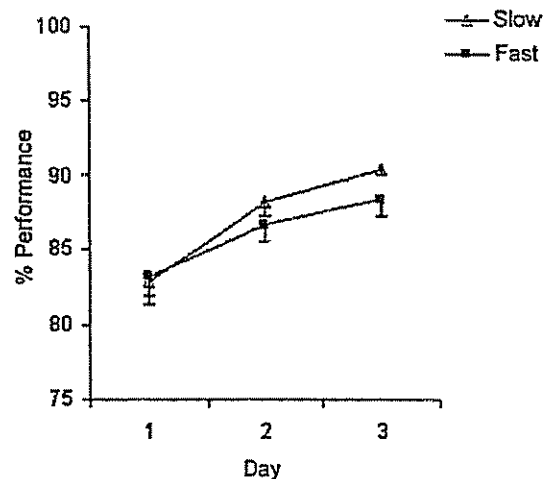


Figure 2. Performance by time condition.

performance when tested under time constraints on the 3rd day, even though they had not previously been exposed to the fast condition. As an index of participant performance, the average performance achieved by fast-condition participants on the 3rd day was close to the performance predicted by the T decision heuristic (122, or 87.9%), whereas the average performance of slow-condition participants on the 3rd day was close to the performance predicted by the TV heuristic (99, or 90.2%).

Figure 3 shows the relationship between learning and Raven score over the 3-day period. The top panel shows the main effect of Raven score. For analytical purposes, participants were classified as high or low Raven according to their mean Raven score (54), but all statistical analyses were performed by using the results of the Raven test without the binary variable. Participants with a Raven score above 54 were

classified as high Raven ( $N = 16$ , mean = 57.0,  $SD = 1.75$ ), and participants with a score below or equal to 54 were classified as low Raven ( $N = 17$ , mean = 50.4,  $SD = 3.18$ ). Note that the *low Raven* and *high Raven* designations refer to the population in this experiment only and do not refer to the overall population identified by Raven et al. (1977). In the fast-condition group, 7 participants were high Raven and 7 were low Raven. In the slow-condition group, 9 participants were high Raven and 10 were low Raven. High-Raven participants performed significantly better than low-Raven participants throughout the experiment, regardless of condition.

The bottom panels in Figure 3 show the Day  $\times$  Time  $\times$  Raven score interaction. In the fast-condition group, the difference in performance between low- and high-Raven participants increased during the practice sessions and continued to increase after transfer (i.e., under testing

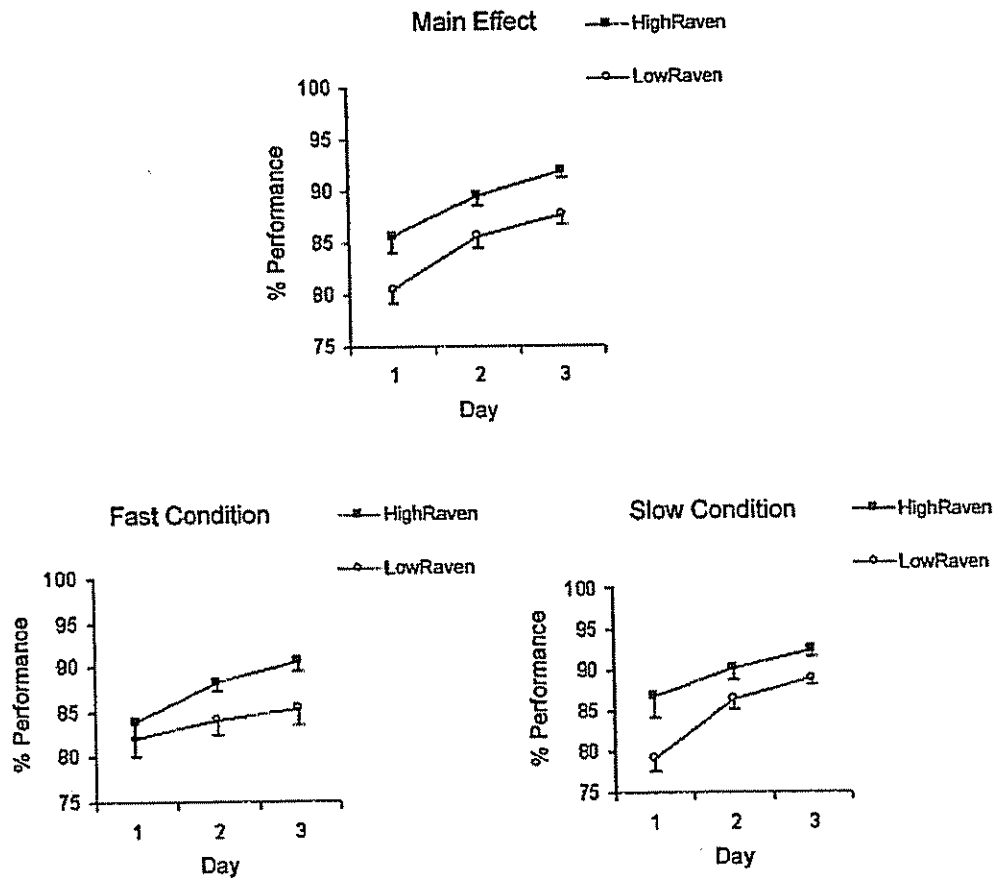


Figure 3. Performance. Top panel: Main effect of Raven score. Bottom panels: Day  $\times$  Time  $\times$  Raven score interaction.



conditions), whereas in the slow-condition group, the difference in performance between the low- and high-Raven participants decreased with each day. In the fast-condition group, high-Raven individuals outperformed low-Raven participants by a considerable margin overall. Conversely, the absence of time constraints on the practice days greatly benefited the performance of low-Raven participants in the slow-condition group.

To eliminate any possible confounding effects attributable to task practice, a second set of analyses was conducted to compare the 10 trials performed by the slow-condition participants with the first 10 trials performed by the fast-condition participants. The analyses revealed the same significant effects previously identified by the per day analyses: a main effect of Raven score,  $F(1, 27) = 10.3, p = .003$ , significantly different learning between conditions,  $F(9, 243) = 3.34, p = .001$ , and significant learning between conditions modulated by Raven score,  $F(9, 243) = 3.18, p = .001$ . Again, the detrimental effects of time pressure on learning were clear when comparing the same number of trials for the slow and fast conditions, and cognitive ability, as measured by the Raven test, was determinant of performance.

The fit between participants' actual decisions and the decisions as predicted by each of the three heuristics (T, TV, TVC) was analyzed in the same way as the performance measures. The average of the fit for all decisions was calculated for each participant and per day (i.e., each participant had three fit values corresponding to

each of the three heuristics, for each decision) and was used as repeated measures in the analyses. Raven score again was used as a covariate. The average fit calculations for all three heuristics are shown in Table 2. The time constraints had a main effect on the fit of the T and the TVC heuristics. Also, there were significant interactions of time constraints with day and with the interaction of day and Raven. Because the fit to heuristics per trial revealed results similar to those calculated per day, only the results per day are shown in the table.

The graphs presented in Figure 4 describe only the T heuristic, but the results for the TV and TVC heuristics were similar to the T heuristic results. The top panel of Figure 4 indicates that the decisions made by participants in the slow and fast conditions fit the T heuristic differently. The decisions of slow-condition participants showed a decrease in fit to the T heuristic over the 2 practice days, whereas decisions by the fast-condition participants fit the T heuristic similarly over time.

As observed in the performance comparisons, Raven score had a significant effect on the fit to the heuristics. The bottom panels of Figure 4 depict how heuristic fit was affected by time constraints and the cognitive abilities of the participants. Under time constraints (i.e., the fast condition), decisions made by low-Raven participants during practice and after transfer increasingly fit with the decisions predicted by the T heuristic. The opposite pattern was observed upon analysis of the decisions made by

TABLE 2: Repeated-Measures ANOVAs for the Three Decision Heuristics

	Time			Time Volume			Time Volume Chain		
	DF	F	p	DF	F	p	DF	F	p
Within Subjects									
Day	2, 58	1.35	.26	2, 58	1.682	.19	2, 58	2.05	.13
Day × Time	2, 58	5.33	.007**	2, 58	4.36	.017*	2, 58	1.57	.21
Day × Raven	2, 58	1.66	.19	2, 58	2.05	.13	2, 58	2.22	.11
Day × Time × Raven	2, 58	4.90	.011*	2, 58	4.04	.023*	2, 58	1.41	.255
Between Subjects									
Time	1, 29	4.57	.041*	1, 29	3.831	.060	1, 29	5.55	.025*
Raven	1, 29	0.007	.93	1, 29	0.369	.549	1, 29	2.12	.15
Time × Raven	1, 29	4.98	.033*	1, 29	4.08	.053	1, 29	5.97	.021*

\*p < .05. \*\*p < .01.

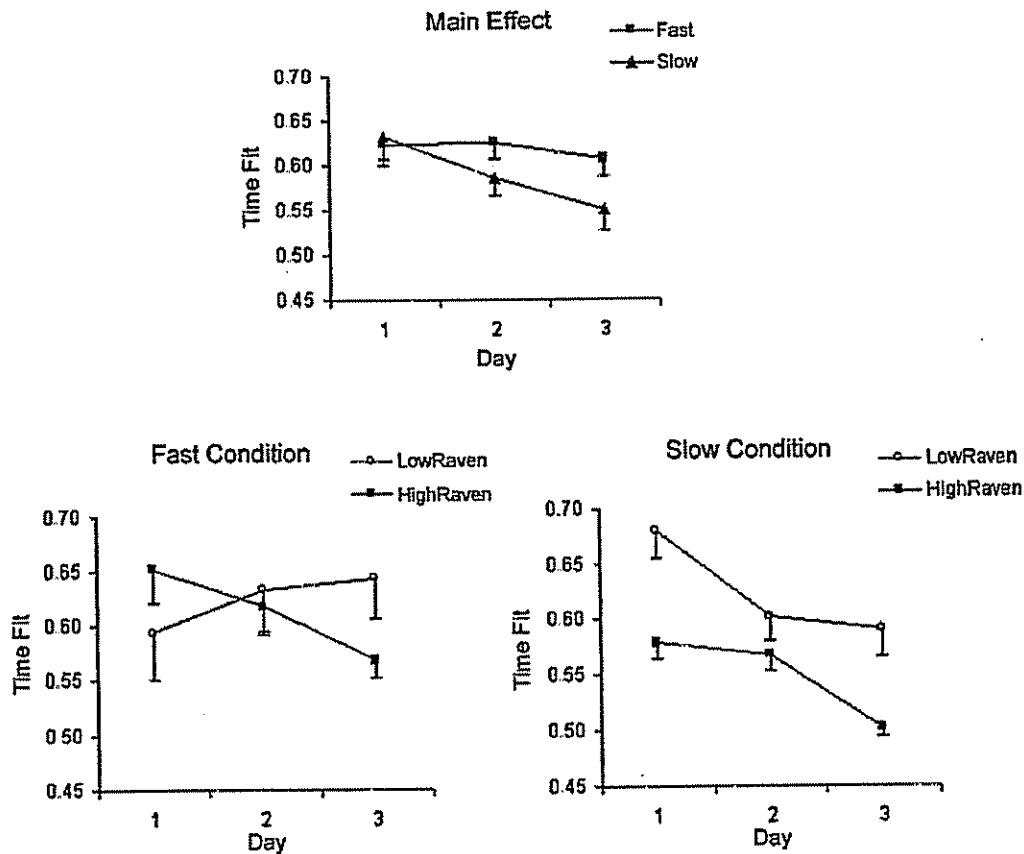


Figure 4. Average fit to the time heuristic. Top panel: Main effect of time condition. Bottom panels: Day  $\times$  Time  $\times$  Raven score interaction.

high-Raven participants. In contrast, under the slow condition the fit between the decisions of both high- and low-Raven participants and the decisions predicted by all three heuristics decreased over time. The decisions made by high-Raven participants in the slow condition fit the T heuristic less closely than did those of low-Raven participants, particularly on the 3rd day (i.e., after transfer).

Regression analysis was performed using the heuristic fit, the Raven score, and the standard deviation of the average fit as measures of variability in following a heuristic. This regression analysis demonstrated that the better the fit to a heuristic, the worse the performance. This result held for all three heuristics: for T,  $\beta = .205$ ,  $t = 4.387$ ,  $p < .001$ ; for TV,  $\beta = .188$ ,  $t = 3.866$ ,  $p < .001$ ; and for TCV,  $\beta = .141$ ,  $t = 3.997$ ,  $p < .001$ . (Positive coefficients indicate that an increase in the percentage fit to the heuristic will increase the number of missed gallons.) The Ra-

ven score and the standard deviation of the heuristic fit also were regressed positively to performance. Higher Raven scores indicated improved performance: for T,  $\beta = -.105$ ,  $t = -2.754$ ,  $p = .007$ ; for TV,  $\beta = -.104$ ,  $t = -2.659$ ,  $p = .009$ ; and for TCV,  $\beta = -.085$ ,  $t = -2.339$ ,  $p = .021$ . Additionally, a higher standard deviation of the average heuristic fit led to better performance: for T,  $\beta = -.148$ ,  $t = -3.773$ ,  $p < .001$ ; for TV,  $\beta = -.176$ ,  $t = -3.989$ ,  $p < .001$ ; and for TCV,  $\beta = -.210$ ,  $t = -6.208$ ,  $p < .001$ . The positive performance effects observed with a decreasing fit to the heuristics increased from the T to the TV to the TVC heuristic (from  $-.148$  to  $-.210$ ).

## DISCUSSION

The most intriguing result generated by this study is the finding that mere repetition of a task by a decision maker under time constraints does not translate into better task performance.

Participants who completed more trials under severe time constraints exhibited inferior performance in comparison with participants who completed fewer trials and who had more time for reflective decision making within each trial. Furthermore, less stringent time constraints during practice sessions resulted in better performance by participants when they subsequently were tested under severe time constraints.

Additional analysis of these experimental results indicates two possible explanations for this key finding. First, the detrimental effects of time constraints on participant performance may be attributable to the participants' inability to solve novel situations and adapt to new changing conditions (i.e., to the participants' finite cognitive abilities). The study showed an interaction of time constraints and cognitive abilities, indicating that less stringent time constraints during the practice sessions particularly benefited the low-Raven participants. These participants' performance also transferred well when they were tested under high time constraints. The psychology literature suggests that high Raven scores reflect an individual's ability to recognize abstract relationships and to manage a larger set of goals in the working memory than can individuals with low Raven scores. These abilities are essential to learning in DDM, which involves situations that change continuously and the appearance of several alternatives in close temporal proximity. Dynamic decision makers must internalize the constantly evolving progress of the DDM task while exploring possible new paths by which they might complete the task.

The analyses of how well participants' actual decisions corresponded with the decisions predicted by heuristic models offer a second possible explanation of the findings in this study. With increased task practice, low time constraints were associated with a poorer heuristic fit, independent of the individual's cognitive abilities. Conversely, the heuristic fit improved with task practice only for individuals with low cognitive abilities who were under severe time constraints. This finding suggests that more lengthy trials and availability of cognitive abilities helped individuals acquire more complex and useful knowledge that they subsequently were able to use when tested under time constraints. Indi-

viduals with low cognitive abilities who were under time constraints were perhaps unable to acquire more complex and context-based knowledge, relying more on simple heuristics.

The type of knowledge attained through practice in a dynamic task and the process by which individuals acquire such knowledge is hypothesized to be instance-based decision making (Gonzalez et al., 2003). Although instance-based knowledge acquisition was not demonstrated by the experiment described here, one might reasonably hypothesize that individuals under low time constraints and with high cognitive abilities are able to acquire context-based knowledge that they can later use if asked to perform under time constraints. The variability in following heuristics appears important to better learning, as this study suggested a positive correlation. According to the instance-based model, it is possible that this variability leads to the creation of a diverse repertoire of exemplars that would support adaptability to unexpected conditions. The absence of stringent time constraints may permit individuals to evaluate more alternatives and thus create a more diverse repertoire of exemplars.

This interpretation regarding heuristics needs to be corroborated in future research. The data reported here were obtained by comparing the participants' behavior with that behavior dictated by specific heuristics. However, individuals were not asked to use any heuristics, nor were they requested to report their use of heuristics. Thus these data do not confirm the acquisition of instance-based knowledge. Further, the analyses of heuristics in this paper involved a particular set of possible simple heuristics derived from verbal protocols, but it is likely that individuals use other, nontested heuristics for which the results don't necessarily generalize.

#### ACKNOWLEDGMENTS

I would like to acknowledge the support of Dr. Javier Lerch during the development of this experiment. Also, I would like to thank Dr. Don Harter for his assistance with data collection and the anonymous reviewers and Nancy Cook for excellent and helpful reviews. This research was supported by the Advanced Decision Architectures Collaborative Technology

Alliance sponsored by the U.S. Army Research Laboratory (DAAD19-01-2-0009) and the Multidisciplinary University Research Initiative Program (MURI), Grant Number N00014-01-10677.

## REFERENCES

- Benson, L., & Beach, L. R. (1996). The effects of time constraints on the prechoice screening of decision options. *Organizational Behavior and Human Decision Processes*, 67(2), 222–228.
- Brehmer, B. (1990). Strategies in real-time, dynamic decision making. In R. M. Hogarth (Ed.), *Insights in decision making* (pp. 262–279). Chicago: University of Chicago Press.
- Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, 9(2–3), 171–184.
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven progressive matrices test. *Psychological Review*, 97, 404–431.
- Dreyfus, H. L. (1997). Intuitive, deliberative, and calculative models of expert performance. In C. E. Zsombok & G. Klein (Eds.), *Naturalistic decision making* (pp. 17–28). Mahwah, NJ: Erlbaum.
- Edland, A., & Svenson, O. (1993). Judgment and decision making under time pressure: Studies and findings. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 27–38). New York: Plenum.
- Edwards, W. (1962). Dynamic decision theory and probabilistic information processing. *Human Factors*, 4, 59–73.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, 128, 309–331.
- Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
- Gonzalez, C. (2003). Verbal protocols in real-time dynamic decision making. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting* (pp. 293–296). Santa Monica, CA: Human Factors and Ergonomics Society.
- Gonzalez, C. (in press). The relationship between task workload and cognitive abilities in dynamic decision making. *Human Factors*.
- Gonzalez, C., Lerch, F. J., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27(4), 591–635.
- Hogarth, R. M., & Spyros, M. (1981). Forecasting and planning: An evaluation. *Management Science*, 27(2), 115–138.
- Kerstholt, J. H. (1994). The effect of time pressure on decision-making behaviour in a dynamic task environment. *Acta Psychologica*, 86, 89–104.
- Kerstholt, J. H. (1996). The effect of information costs on strategy selection in dynamic tasks. *Acta Psychologica*, 94, 273–290.
- Kerstholt, J. H., & Raaijmakers, J. G. W. (1997). Decision making in dynamic task environments. In W. R. Crozier & O. Svenson (Eds.), *Decision making: Cognitive models and explanations* (pp. 205–217). London: Routledge.
- Lerch, F. J., Ballou, 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 automatization. *Psychological Review*, 95, 492–527.
- Maule, A. J., & Edland, A. C. (1997). The effects of time pressure on human judgment and decision making. In R. Ranyard, W. R. Crozier, & O. Svenson (Eds.), *Decision making: Cognitive models and explanations* (pp. 189–204). London: Routledge.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.
- Rastegary, H., & Landy, F. J. (1993). The interactions among time urgency, uncertainty and time pressure. In O. Svenson & A. J. Maule (Eds.), *Time pressure and stress in human judgment and decision making* (pp. 217–239). New York: Plenum.
- Raven, J. C., Court, J. H., & Raven, J. (1977). *Standard progressive matrices*. London: Lewis.
- Svenson, O., & Maule, A. J. (1993). *Time pressure and stress in human judgment and decision making*. New York: Plenum.

Cleotilde Gonzalez is an assistant professor in the Social and Decision Sciences Department and director of the Dynamic Decision Making Laboratory at Carnegie Mellon University. She received her Ph.D. in management information systems in 1995 from Texas Tech University.

Date received: March 17, 2003

Date accepted: January 26, 2004