VERBAL PROTOCOLS IN REAL-TIME DYNAMIC DECISION-MAKING

Cleotilde Gonzalez
Social and Decision Sciences Department
Carnegie Mellon University
Pittsburgh, PA
conzalez@andrew.cmu.edu

This study presents the results from the analyses of verbal protocols elicited from inexperienced and experienced participants of a real-time, Dynamic Decision-Making (DDM) task. This research intends to complement a series of studies performed in DDM environments analyzing the cognitive structures and processes involved in learning in DDM. Results show that inexperienced and experienced participants differ in several ways: in the way they distribute attention to different parts of the system, in their awareness of the relationship of the attributes involved in the decision making process, and in their coordination to make decisions in real time. These results have been used to support the refinement of a cognitive model developed to explain how people learn in DDM tasks.

INTRODUCTION

Real-time, Dynamic Decision Making (DDM) has three main characteristics: a) multiple and interdependent decisions; b) the environment changes because of exogenous events and because of prior decisions; and c) the pacing of decisions is dictated by the task rather than by the decision maker (Brehmer, 1990). Verbalizations of the thought process are often used to study expertise and decision making in real-world complex tasks (Hoffman, Shadbolt, Burton, & Klein, 1995). These methodologies are also frequently used as the basis to build cognitive models. Development of cognitive models often requires detailed information regarding attention and cognitive processes in the task. In particular, it is necessary to know how and why participants select alternatives, in which order, and how they evaluate and judge them. One way towards understanding these detailed cognitive processes is the collection of verbal protocols.

Our research investigates learning in dynamic decision making situations and we have developed cognitive models of the learning process (Gonzalez, Lerch, & Lebiere, 2003). Often, protocols are elicited from experts, and only rarely research looks at verbalizations from people with different experience. The work initiated by Herbert Simon on chess playing suggests that experts and novices follow different mental processes. Experts search very selectively, using environmental cues to guide their attention while novices engage in more exhaustive search (Simon & Gobet, 1996).

This paper reports a study on the collection and analysis of verbal protocols from individuals with different experience in a DDM task. These protocols can inform modelers of the search, judgment and choice strategies development with experience.

METHOD

The Dynamic Decision Making task

Water Production Plant (WPP) simulation is an abstraction of a resource management task occurring in a real-world organization. WPP simulates a water distribution system with multiple deadlines for alternative tanks in the system. Decision makers have to decide when to activate or deactivate pumps associated with different tanks, given a restriction in the number of pumps working at any given time. Figure 1 shows a screenshot of the simulation, and a detailed description of the task can be found in other publications (Gonzalez et al., 2003).

WPP is highly dynamic because water may arrive into a tank at any time, and the level of water in each tank depends on prior decisions (i.e., the pumps that were activated or deactivated by the decision maker in the past). The task is also real-time because the simulation clock is running continuously while participants make their decisions. One measure of performance in this task is the number of gallons of water that were not pumped in time, therefore the higher the number of gallons missed, the worse the performance.

Participants

A total of four subjects (two men and two women) participated in this study. Each participant was paid $10 per hour. Participants ran the WPP simulation at a rate of 8 minutes per trial, simulating 8 hours of water processing.

Experimental design

A heterogeneous pair of individuals was selected to participate in the inexperienced condition and two other

individuals participated in the experienced condition. Figure 2 presents the activities of both inexperienced and experienced participants. In the first hour, all participants completed a training session in the WPP simulation.

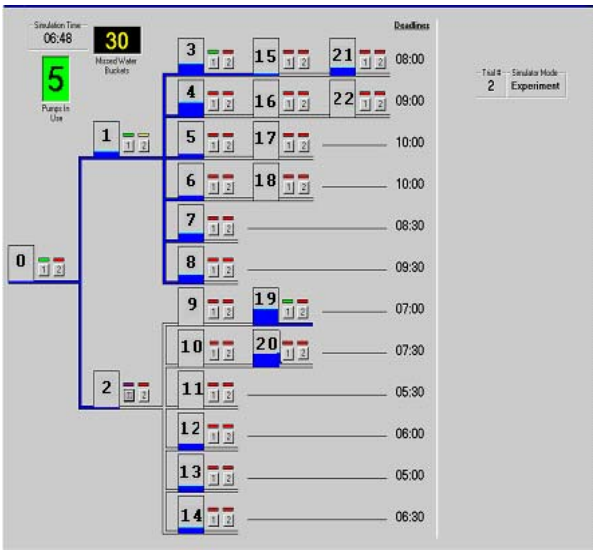


Figure 1. Layout of the Pipes Task

Then, they were trained to think aloud and forced to verbalize their thoughts while doing simple practice problems such as multiplication of two numbers and a simple imagery task.

In the second hour, participants in the inexperienced group performed the WPP once while they were also trained for verbal protocols in this task. Then, they performed the simulation one more time before protocols were collected. Participants assigned to the experienced condition performed the task 18 times during 3 consecutive days. On the third day, participants were trained for verbal protocols and performed the simulation one more time before protocols were collected. In total, participants in the inexperienced condition spent 16 minutes in the simulation while experienced individuals spent 152 minutes. The length of the verbal protocol collection time was the same for both inexperienced and experienced subjects: 8 minutes. Participants were not given any strategies on how to perform better in the task.

	Hour 1	Hour 2	Hours 3&4	Hours 5&6	Total Time
Inexperienced	Training in simulation	1 trial, protocol training,1 trial protocol collection			16 min.
Experienced	Training in simulation	5 trials	10 trials	3 trials, Protocol training, 1 trial, protocol collection	152 min.

Figure 2. Experimental conditions

Protocol Collection

Due to the characteristics of WPP (highly dynamic and interrelated decisions on real-time), we decided to collect retrospective rather than concurrent protocols. It has been found that when participants verbalize the bases of their decisions in dynamic decision tasks, they perform significantly worse than when they are silent (Dickson, McLennan, & Omodei, 2000). The protocols collected are not the commonly known as retrospective reports (Ericsson & Simon, 1993). For the collection of our protocols, we used the "replay" feature of the WPP simulation. This allows us to see every single action performed by the subject at the same speed those actions were executed. The advantages of using a replay to collect the protocol are that, subjects would not feel overloaded, because they don't have to execute the task while talking, and yet easily recall the actions taken during the simulation. We replayed the simulation right after the participants in both conditions performed the last trial. Both videotape and audiotape were used to collect the protocols.

Protocol Analysis

Protocols were first transcribed verbatim from the tape recording. Next, protocols were segmented into separate lines to identify specific ideas the participants had verbalized. The segmented protocols were encoded and analyzed using the Protocol Analyst's Workbench (PAW) software (Fisher, 1991).

PAW is an interactive protocol analysis system that supports the encoding and analysis of verbal protocols. The tool allows encoding the protocols by separating and organizing different actions and objects. For example, in the segmented protocol we have the following phrase:

"I think I wanted to devote resources to the 5:00 o'clock because that is the first deadline."

This is translated into action (Explain) involving several objects (Tank, Pump, Deadline) as shown in line 5 of Figure 3. Each segment in PAW represents a statement with an action and a set of descriptors. PAW can create a process model and analyze the descriptor values used in the protocol, trace cycles, present the statements and types of cycles in the protocol, compare the values of descriptors for each cycle type, and compare the process models for different processes and protocols. These tools were used in the comparative analysis of protocols from inexperienced and experienced participants, as presented next.

RESULTS

Inexperienced participants reached an average score in WPP of 224.5 (gallons of water) while experienced participants reached an average score (out of 19 trials) of 114.6, and a minimum score of 52 gallons. Again, the score refers to the number of gallons of water missed, with higher numbers indicating poorer performance.

```
PROTOCOL:
-------------- <1> --------------
1     Tank      pumps
Activate        C          2
-------------- <1> --------------
2     Tank      pumps
Activate        B          1
-------------- <1> --------------
3     What
Start Simulation
      Tank      pumps
Activate        B1.1       2
-------------- <1> --------------
4     Tank      pumps
Activate        C1.5       1
-------------- <1> --------------
5     Tank      Ref        Order
Explain         C1.5       £4          First deadline
-------------- <1> --------------
6     Tank      pumps
Activate        C1.3       1
-------------- <1> --------------
```

Figure 3. A segment of the PAW protocol

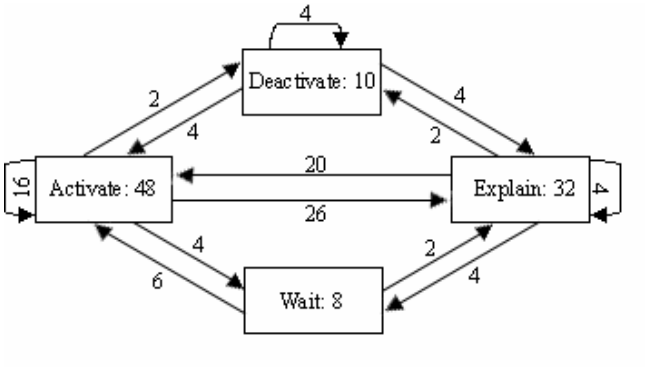
Figure 4 presents the process model for inexperienced and experienced participants shown as a network. To construct the networks, we used the transition matrix produced by PAW, indicating the actions, orders, and frequencies with which those actions appear in the protocols. Each rectangle in the network represents an action, the lines represent the paths from/to a node from/to another node. The numbers in the paths reflect the total number of times a path was traversed. The curved arrows represent the number of times a node is followed by itself. The numbers inside the nodes represent the total number of times an action was verbally expressed in the protocol.

Our analysis showed that inexperienced people had more trouble verbalizing their actions than experienced participants. Inexperienced individuals made 10 comments that could not be classified as Actions-Descriptors compared to 3 from the experienced. The total number of verbalizations was slightly higher for experienced compared to inexperienced (108 statements for experienced and 98 for experienced).

The most active nodes for both conditions are: Activate and Explain. The process model indicates that, for inexperienced participants 10% of the path traversals (10/98) were between Activate and Wait activities, while 24% were for the condition. The action "Wait" was verbalized twice as often in the experienced compared to the inexperienced condition. One main difference between these two conditions is that experienced participants waited more after and before activating pumps, compared to inexperienced participants.

Of particular importance to this study were the descriptors related to the time (i.e., deadlines), the amount of water in the tanks, and the position of the tanks in the tree structure (called "chain"). A general analysis of the verbalization of "Tank" and "Deadline" in the action "Activate" revealed that inexperienced participants concentrated on the tanks located in the lower part of the screen (deadlines 5:00 to 7:30) while experienced participants distributed their verbalization almost equally among the tanks across the deadlines, with just slightly more attention to the upper part of the screen (i.e., tank B1.6, 9:30). Inexperienced participants verbalized the earlier deadlines only, while experienced took note of several other deadlines distributed in the lower and upper part of the screen.

Inexperienced:



Experienced:

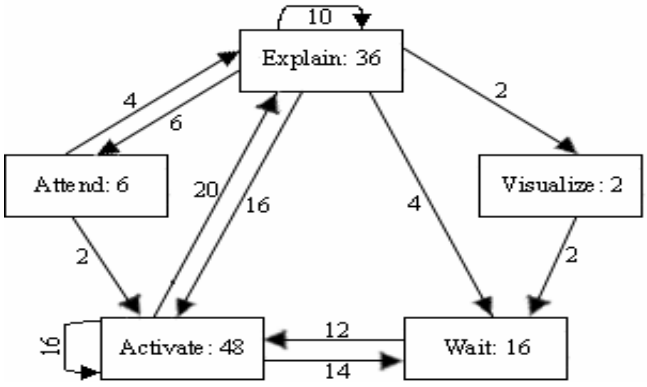


Figure 4. Process Models for inexperienced and experienced participants.

About the amount of water in the tanks, inexperienced and experienced participants were similar in their verbalizations. Individuals in the two conditions agreed on many levels to describe the amounts of water, and used about the same categories to describe the amounts of water in the tanks. Also they were aware of the structure and depth of the tank tree. However, compared to inexperienced, experienced were more aware of the relationship between the depth of the tank, water, and the time required to distribute water to the destination.

A comparison of trace cycles was performed for Explain, Activate and Wait. Results for types of cycles in both, inexperienced and experienced conditions are very different. Only a small percentage of the types of cycles are shared in both inexperienced and experienced conditions: 37.50% for Explain, 42.86%, for Activate and 25.00% for Wait. However, the similarity of the cycles shared by experienced and inexperienced participants is high: 64.71% for Explain, 83.33% for Activate and 42.86% for Wait. This may indicate that those cycles not shared by inexperienced and experienced participants are responsible for differences in performance. For example, one of the participants in the inexperienced condition performed the following cycle:

Wait->Explain->Deactivate->Explain->Deactivate->Activate->Deactivate->Activate->Explain->Activate->Explain->Activate->Explain->Wait

While a participant in the experienced condition performed the following cycle:

Wait->Activate->Explain->Activate->Wait

These cycles suggest that inexperienced participants performed many actions before waiting again, while experienced participants performed fewer actions. Among the actions performed by participants in the inexperienced group, there are several Activate and Deactivate actions that, as correctly noted, cause "switching costs".

CONCLUSIONS

Results show that the behavior of people with and without experience in a dynamic task differs in several ways. First, experienced people are more aware of where to attend and how to give prognosis of a situation than inexperienced participants. Second, experienced participants learned to wait to make decisions and try to analyze and explain their actions more than did inexperienced participants. This is important in DDM environments, where many decisions may imply additional costs and therefore affect performance. Third, experienced people seem to divide their attention equally throughout the system compared to inexperienced participants that concentrated only on some parts of the system. In DDM tasks it is very important to evaluate and be aware of all the activities at any moment of time. Fourth, people with more experience seem to be more aware of the relationship of the different variables involved in making a decision, while less experienced people seem to focus on one aspect at a time while making decisions. Finally, experienced individuals know when to wait, it appears that they know when to act at the right time.

DDM tasks present several challenges for the collection and analysis of verbal protocols. The dynamic and highly complex nature of these tasks inhibits concurrent verbalization of cognitive processes. The collection of protocols with replay is helpful to recall all the actions while verbalizing, without the need of executing the task at the same time. However, as expressed and observed from the subjects reported in this study, replaying the simulation made them evaluate their previous decisions rather than just recall what they thought in the past.

Results from these verbal protocols helped us refine the processes implemented in a cognitive model of learning in dynamic decision making. A theory of how people use their past experience to make decisions is in development, and it helps explain how inexperienced people differ and acquire the processing experience found with these protocols (Gonzalez, Lerch, & Lebiere, 2003).

ACKNOWLEDGEMENTS

This research was supported by Advanced Decision Architectures Collaborative Technology Alliance sponsored by the U.S. Army Research Laboratory (DAAD19-01-2-0009), and the Office of Naval research (N00014-96-1-0322).

REFERENCES

1. Brehmer, B. (1990). Strategies in real-time, dynamic decision making. In R. M. Hogarth (Ed.), *Insights in Decision Making* (pp. 262-279): University of Chicago Press.
2. Dickson, J., & McLennan, J, & Omodei, M. M. (2000). Effects of concurrent verbalization on a time pressured dynamic decision task. *Journal of General Psychology*, *127*, 217-228.
3. Ericsson K. A. and Simon H. A. (1993). *Protocol Analysis: Verbal Reports as Data*. Cambridge: The MIT Press.
4. Fisher, C. (1991). *Protocol Analyst's Workbench: Design and Evaluation of Computer-Aided Protocol Analysis.* Thesis: Pittsburgh, PA. Carnegie Mellon University.
5. Gonzalez, C., Lerch, F. J., & Lebiere, C. (2003). Instance-Based Learning in Dynamic Decision Making. *Cognitive Science. 27,4*.
6. Hoffman, R. R., Shadbolt, N. R., Burton, A. M., & Klein, G. A. (1995). Eliciting knowledge from experienced: A methodological analysis. *Organizational Behavior and Human Decision Processes, 62*, 129-158.
7. Simon, H. A., & Gobet, F. (1996). Templates in Chess Memory: A Mechanism for Recalling Several Boards. *Cognitive Science, 31*, 1-40.