Carnegie Mellon University
Research Showcase @ CMU

Department of Social and Decision Sciences | Dietrich College of Humanities and Social Sciences

3-3-2005

Task Workload and Cognitive Abilities in Dynamic Decision Making

Cleotilde Gonzalez
Carnegie Mellon University, conzalez@andrew.cmu.edu

Follow this and additional works at: http://repository.cmu.edu/sds

Published In
.

This Article is brought to you for free and open access by the Dietrich College of Humanities and Social Sciences at Research Showcase @ CMU. It has been accepted for inclusion in Department of Social and Decision Sciences by an authorized administrator of Research Showcase @ CMU. For more information, please contact research-showcase@andrew.cmu.edu.

Task Workload and Cognitive Abilities in Dynamic Decision Making

Cleotilde Gonzalez, Carnegie Mellon University, Pittsburgh, Pennsylvania

Researchers often treat workload as a task-dependent variable. To evaluate the effect of workload on individuals' performance, researchers commonly use several methods, such as varying the complexity or number of tasks that test participants are asked to handle or placing individuals under time constraints. Only rarely have researchers investigated workload as a variable dependent on individuals' cognitive abilities. This study investigated workload during dynamic decision making in terms of its dependence on both task workload and cognitive abilities. The findings demonstrate detrimental effects of both high task workload and low cognitive abilities. Further, the results show that high workload is more detrimental in individuals with low cognitive abilities than in individuals with high cognitive abilities. Potential applications of this research include the design of new workload studies and new training protocols in which psychometric tests are used.

INTRODUCTION

Dynamic decisions are real-time decisions that are interdependent and highly constrained by the decision-making environment (Edwards, 1962). For example, many large manufacturing and distribution systems store and disseminate information in real time about the status of the objects within the system. Decision makers then can use this information to alter the system as events unfold. Despite significant advances in information technology, the high information load generated by such dynamic environments continues to pose problems. For example, plane and automobile accidents are more likely to occur when the involved decision makers (i.e., pilots and drivers, respectively) are under heavy workloads.

Excessive cognitive workload is generated when the satisfactory performance of a task demands more resources from the operator than are available at any given time. Although a wealth of research has been performed to evaluate the effects of task demands on human performance, little attention has been paid to the cognitive resources available or to the relationship between task demands and cognitive abilities as individuals acquire experience in a task. The focus of the present study was the relationship between human cognitive abilities and workload in a complex, dynamic decision-making (DDM) task. The experiments reported here enabled the manipulation of task workload experienced by participants and the assessment of participants' general intellectual ability in order to test the hypothesis that the effect of task workload on DDM depends on cognitive ability.

A brief summary of research pertaining to the effects of cognitive workload on DDM appears in the following section, followed by the methods that describe the experiment and measures of cognitive abilities. Next, the results from collected data are reported. Finally, the results are discussed and their implications are presented.

COGNITIVE WORKLOAD AND DYNAMIC DECISION MAKING

Researchers can manipulate task complexity in many ways, such as by altering the number of elements that demand a participant's attention, the number of intermediate decisions required to arrive at a solution, or the display duration (Ackerman, 1988). In general, the manipulation

Address correspondence to Cleotilde Gonzalez, Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213; conzalez@andrew.cmu.edu. **HUMAN FACTORS,** Vol. 47, No. 1, Spring 2005, pp. xxx–xxx. Copyright © 2005, Human Factors and Ergonomics Society. All rights reserved.

Gonzalez,r1.qxd 3/3/05 11:04 AM Page 2

of task complexity affects attention, accuracy, and the time needed to complete a trial (Ackerman, 1988; Ackerman, Beier, & Boyle, 2002; Verwey & Veltman, 1996).

One of the most common workload manipulations involves the assignment of additional tasks to an individual already performing a complex task. Research suggests that the successful performance of multiple concurrent tasks requires the activation of executive control processes (Borkowski & Burke, 1996; Gopher, 1993; Meyer & Kieras, 1997). Damos and Wickens (1980) demonstrated that the mental demands that executive processes and time-sharing skills place on decision makers are distinct from resource demands of the individual tasks and, thus, produce some interference during dual-task training. This dual-task interference usually results in poorer performance of both tasks (often called *dual-task decrement*) when compared with the completion of each task alone. In an attempt to explain why increased workload has detrimental effects on dual-task skill acquisition, researchers have suggested that cognitive resources are shared during the performance of the competing tasks and, therefore, that individuals must possess executive abilities to successfully complete complex tasks (Damos & Wickens, 1980; Gopher & Donchin, 1986; Wickens, 1980, 1987, 1991).

The use of time constraints is another common method of workload manipulation. Some studies have investigated the effects of time constraints on DDM. For example, Kerstholt (1994) observed accelerated information processing with increasing time constraints (i.e., as the speed of the system increased). Other studies have indicated poorer decision-making performance by individuals performing dynamic tasks while under time constraints than by those performing static tasks while under time constraints (Kerstholt, 1995; Payne, Bettman, & Johnson, 1993).

Despite numerous attempts to assess the effects of task demands on human performance, little is known about the role that cognitive abilities play in learning DDM. By studying static tasks, researchers have demonstrated that general cognitive abilities have an increasing effect on task performance as task complexity increases (Carpenter, Just, & Shell, 1990; Kyllonen & Christal, 1990). Thus one might speculate that cognitive abilities also play a key role in complex dynamic systems, especially in individuals' ability to improve DDM performance with practice.

Intriguingly, current research seems to contradict this expectation. Rigas and Brehmer (1999) observed only a low correlation between psychometric intelligence, as measured by the Raven Standard Progressive Matrices Test (Raven, Court, & Raven, 1977), and performance on two dynamic decision tasks. They explained these findings by hypothesizing that the cognitive processes required to complete the Raven test involve simple combinations of basic operations whereas the cognitive processes needed to complete the dynamic tasks in the study were more complex. This rationale calls into question both the ability of psychometric tests to predict real-world performance and the notion that cognitive abilities directly influence DDM. More recently, however, the same researchers reported significant correlations between cognitive abilities, as measured by the Raven Standard Progressive Matrices Test, and DDM performance (Rigas, Carling, & Brehmer, 2002), results that conflict with their earlier findings.

The study reported in this paper is based on the hypothesis that cognitive abilities play a key role in the adequate management of workload by individuals performing complex tasks. Using two common forms of workload manipulation (dual tasks and time constraints), this study tested whether a higher task workload has a greater effect on individuals with low cognitive ability than on those with high cognitive ability.

METHOD

Because control of a dynamic system is acquired only with extended practice (Kerstholt & Raaijmakers, 1997), the experiment was designed to place participants under different kinds of task workload with lengthy and equal practice in a DDM simulation.

Fifty-one students (31 women and 20 men) were recruited from local universities and paid $50 to participate in the study. Participants were assigned to one of three workload conditions: slow, fast, or load. The slow condition was designed to give participants the lowest task workload possible in a DDM simulation. While performing each simulation trial, individuals in

Gonzalez,r1.qxd 3/3/05 11:04 AM Page 3

the slow condition were presented with a set of events that had to be resolved in 24 min (real time). Participants assigned to the fast and load condition were placed under heavier task workloads than were participants in the slow condition. Individuals in the fast condition had to accomplish the same number of events as those in the slow condition but had to do so in one third of the time (i.e., in 8-min rather than 24-min trials). The load condition participants had to complete the same DDM simulation at the same pace as slow condition participants, but they also had to simultaneously perform two additional, independent tasks. Under each of the three conditions, participants performed the same DDM simulation comprising the same number of events.

Participants ran the DDM simulation on 3 consecutive days. The first 2 days, during which participants worked under one of the three workload conditions, constituted the practice phase. On the 3rd day, during the test phase, all participants performed the same DDM simulation at a fast pace for 48 min. During the practice phase, participants in the slow and load condition groups completed two 24-min trials per day (48 min/day) and participants in the fast condition group completed six 8-min trials per day (48 min/day). During the test phase, participants in all groups completed the same number of trials (six) at the same pace (8 min/trial, 48 total min on task). Thus all participants spent the same total amount of time on the task over the 3-day period: 48 min per day for a total of 144 min. This design facilitated an investigation into the relationship between cognitive ability and workload as individuals transferred from one workload condition to another. Before the practice phase, all participants also completed the Raven Standard Progressive Matrices Test (described later) as a measure of their cognitive abilities.

Dynamic Decision-Making Task

All experimental conditions were based on a DDM task called the *Water Purification Plant* (WPP; Gonzalez, Lerch, & Lebiere, 2003). The WPP simulates a water purification system constructed of a series of tanks joined by pipes. A maximum of five pumps can be active at any given time, and the participant needs to select which pumps to open or close to distribute all the water in a series of tanks as various deadlines approach and expire. A screen shot of the WPP simulation is provided in Figure 1. The WPP simulation constitutes a dynamic task for several reasons: Decisions are interconnected because some actions may delay or preclude other decisions; the amount of water in any of the tanks may increase at any time (in response to a preset scenario of water arrival times and locations that is unknown to the users and beyond their control); the level of water in each tank depends on prior decisions (i.e., the user's earlier activation or deactivation of the pumps); and a time delay occurs after the activation or deactivation of any pump (i.e., pump clean-up time). The WPP is a real-time simulation in that the pumps are activated or deactivated by the users while a simulation clock is running.

The WPP simulation requires participants to pump a total of 1080 gallons (4088 L) of water through the series of tanks. Performance in this task is measured by the total number of gallons of water remaining in the system at the end of the simulation. Thus the best performance is zero, and the performance if no action is taken in the system is 1080 gallons. A running counter in the upper left corner of the screen indicates the number of gallons of water left in the system after the expiration of each deadline. For data analysis, the number of missed gallons was converted to the percentage of the total gallons pumped out of the system; therefore performance could range from 0% to 100%. In order to establish a reasonable lower limit for the performance measure, a program called the *random scheduler* was created to run the simulation and make random assignments with no idle time (i.e., never leaving pumps idle). Thirty replications of these random assignments generated a mean performance of 83% with a standard deviation of 2.6%. Accordingly, the lowest reasonable human performance should score in the range of 80%.

The WPP task for individuals in the load condition group was exactly the same as described, except that these participants were also asked to simultaneously and independently perform two additional tasks, labeled *system monitoring* and *communications.* Figure 2 shows the layout of the WPP with these additional tasks.

Gonzalez,r1.qxd 3/3/05 11:04 AM Page 4

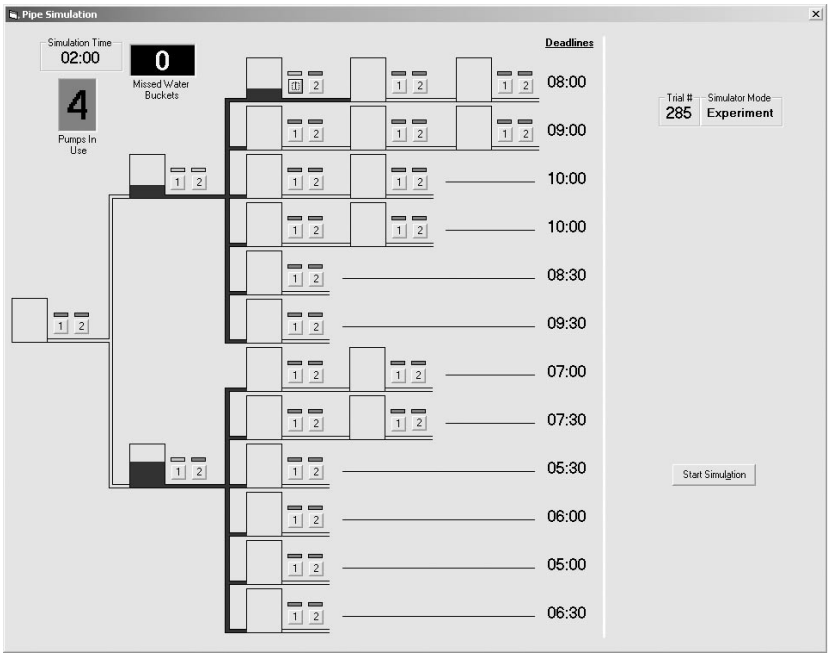

Figure 1. Layout of the WPP task. Water enters from outside the system and moves continuously through the activated pumps from left to right toward the deadlines. The operator decides when to activate and deactivate pumps while the simulation time is running.

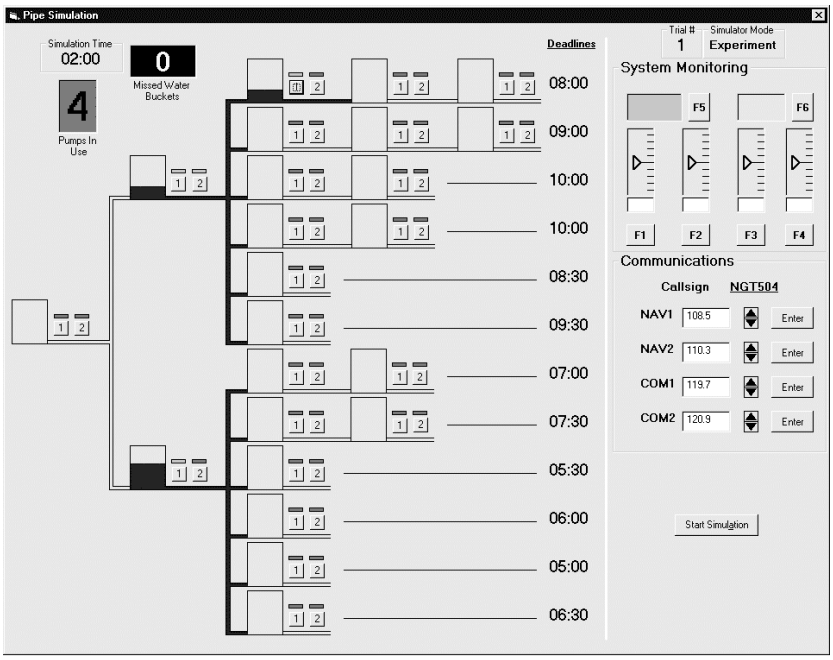

Figure 2. Layout of the WPP task with additional tasks under the load condition. Besides performing the WPP task, operators must simultaneously monitor the gauges on the right of the screen and attend to auditory messages to update the communications settings.

Gonzalez,r1.qxd 3/3/05 11:04 AM Page 5

These two tasks are components of the Multi-Attribute Task Battery developed at the National Aeronautics and Space Administration by Comstock and Arnegard (1992). These two tasks are not integrated with the WPP but, rather, run in parallel with and independently of the WPP. The tasks stop concurrently with the WPP simulation. The system monitoring task requires users to monitor two warning lights (a green light and a red light) and four vertical gauges that report system abnormalities. The communications task requires users to discriminate audio signals and respond to their own call sign (e.g., NGT504) by making frequency changes on the proper navigation or communication radio. The performance measures in these two additional tasks were the percentage of correct responses and the response time. Training for the two additional tasks was separate from the training in the WPP. Before the start of the experiment, participants were asked to pay equal attention to the WPP and the loading tasks during their simultaneous performance.

Cognitive Abilities Measure

The Raven Standard Progressive Matrices Test was used to evaluate the participants' cognitive abilities (Raven et al., 1977). The Raven test is nonverbal and relatively free of cultural bias. Although the use of this measure to evaluate people's ability to perform DDM tasks is unproven (Rigas & Brehmer, 1999), research in psychology suggests that the Raven test is a good indicator of an individual's ability to dynamically manage a large set of goals in working memory and adapt to new situations (Carpenter et al., 1990). Based on the description and analysis of the Raven test that appear in the psychological literature, the decision to use this test to predict DDM performance seems strongly justified.

The Raven test comprises sets of visual analogy problems, each of which presents a pattern in which a figure is missing at the top of the page. Test takers must select the option that best completes the figure by choosing from among eight alternatives arranged at the bottom of the page. The test includes five sets of problems with 12 questions per set, for a total of 60 questions arranged according to degree of difficulty (with more difficult questions presented at the end). The participant's score is the total number of correct answers (possible range: 0–60). This test requires approximately 40 min to complete.

Statistics

The average performance during the practice phase (average performance on the first 2 days of the experiment) and the average performance during the test phase (the average performance on the 3rd day) were calculated by analyzing the WPP data. Statistical analyses were conducted by repeated measures analyses of variance (ANOVAs), and two phases (practice and test) were used as the within-subjects factor. Workload condition was used as a between-subjects factor, and Raven score was used as a covariate.

Results

During the practice phase, the average performance across all three condition groups was 84% with a standard deviation of 5.3% (minimum = 69%, maximum = 94%, standard error = 0.73%) – a reasonable performance level according to the results generated by the random scheduler. During the test phase, the average performance across all three condition groups was 88%, with a 5.1% standard deviation (minimum = 69%, maximum = 95%, standard error = 0.71%). The average Raven score across the three condition groups was 53.6, with a standard deviation of 4.2 (minimum = 45, maximum = 60, standard error = 0.73). While performing the WPP task, participants made an average of 43.6 decisions per trial (minimum = 20, maximum = 76). Linear regression over the total average performance (across all three days) revealed that Raven score was a good covariate (adjusted $R^2 = .283$), $F(1, 50) = 4.26$, $p < .05$.

Two separate analyses using condition as a between-subjects factor and Raven score as a covariate showed that both condition, $F(2, 48) = 4.19$, $p < .05$, and Raven score, $F(1, 49) = 4.17$, $p < .05$, had independent significant effects on performance. A full factorial model based on both the condition and Raven score indicated main effects of both condition, $F(2, 47) = 7.69$, $p < .01$, and Raven score, $F(2, 47) = 10.07$, $p < .01$. To investigate the interaction between these two factors, I performed statistical analyses with a repeated measures model that included the Condition × Raven interaction. Table 1 shows the

Gonzalez,r1.qxd 3/3/05 11:04 AM Page 6

TABLE 1: Overall Effect of Workload Condition and Cognitive Abilities

	Performance			
	DF	*F*	*p*	Partial Eta^2
Within Subjects				
Phase (practice, test)	1, 45	0.02	.89	.00
Phase × Workload Condition	2, 45	4.12	.02*	.16
Phase × Raven	1, 45	0.47	.50	.01
Phase × Workload Condition × Raven	2, 45	3.93	.03*	.15
Between Subjects				
Workload condition	2, 45	0.08	.92	.00
Raven	1, 45	7.79	.01*	.15
Workload Condition × Raven	2, 45	0.15	.86	.01

*$p < .05$.

results, which indicate a main effect of Raven score, $F(1, 45) = 7.79$, $p < .01$, and two significant interactions: Phase × Condition, $F(2, 45) = 4.12$, $p < .05$, and Phase × Condition × Raven Score, $F(2, 45) = 3.93$, $p < .05$.

These findings indicate that performance in the practice and test phases varied with condition and with Raven score. Figure 3 shows the average performance in the practice and test phases by condition. Repeated measures analyses by condition were performed to evaluate the effect of the phase and the Raven score. The results revealed that only those individuals who practiced under the slow condition (i.e., low workload) improved their performance significantly when subsequently placed under high time constraints during the test phase, $F(1, 17) = 5.72$, $p < .05$. The performance of individuals who practiced under the fast condition did not improve significantly on the 3rd day as compared with their performance on the first 2 days, $F(1, 12) = 1.16$, *ns;* similar results were observed in the load condition group, $F(1, 16) = 1.35$, *ns.* Finally, in the slow condition there was a significant effect of Raven score, $F(1, 17) = 13.51$, $p < .01$, and a significant interaction between phase and Raven score, $F(1, 17) = 3.79$, $p < .05$.

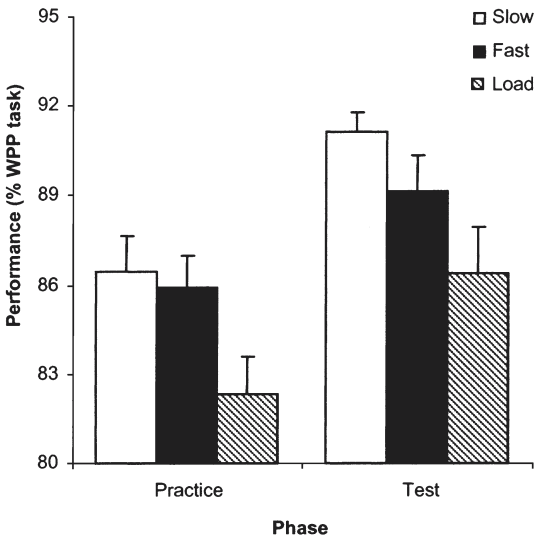

Figure 3. Average performance in the practice and test phases per condition. Error bars represent 95% confidence intervals based on the pooled *MSE.*

Three models involving every possible paired combination of conditions (i.e., slow and fast, slow and load, and fast and load) were tested to further investigate the interactions described previously. Table 2 shows the results, which indicate a significant main effect of Raven score, a Phase × Condition interaction, and a Phase × Condition × Raven Score interaction only for the models that included the slow condition. That is, in the two phases of the experiment, participants' performance in the two groups with high workload (fast and load) differed from that of the group with low workload (slow), but there was no difference in performance between the two high-workload groups.

Participants were classified according to their mean Raven scores as low- or high-Raven individuals. This classification was done only for the purposes of illustrating the Raven Score × Phase interaction and the main effect of Raven score. The sample was not divided, and no analyses were conducted with a split sample; rather,

Gonzalez,r1.qxd 3/3/05 11:04 AM Page 7

TABLE 2: Performance Differences Among Cognitive Load Groups

	Slow vs. Fast			Slow vs. Load			Fast vs. Load		
	$F(1, 29)$	p	Partial Eta^2	$F(1, 33)$	p	Partial Eta^2	$F(1, 28)$	p	Partial Eta^2
Within Subjects									
Phase	1.66	.21	.05	0.06	.81	.00	2.41	.13	.08
Phase × Condition	5.96	.02*	.17	5.10	.03*	.13	0.48	.50	.02
Phase × Raven	0.48	.49	.02	0.03	.87	.00	3.88	.06	.12
Phase × Condition × Raven	5.50	.03*	.16	5.04	.03*	.13	0.51	.48	.02
Between Subjects									
Condition	0.27	.61	.01	0.05	.83	.00	0.01	.94	.00
Raven	14.33	.00**	.33	4.74	.04*	.13	2.54	.12	.08
Condition × Raven	0.37	.55	.01	0.19	.67	.01	0.00	.95	.00

*$p < .05$, **$p < .01$.

statistics were calculated on the whole sample, using the continuous variable of the Raven test score. Referring back to the interactions described previously, in the slow condition improvement in performance was greater for low-Raven individuals than for high-Raven individuals. The left panel in Figure 4 shows that low-Raven participants in the slow condition group greatly benefited from practicing the simulation at the slower pace (i.e., many of them performed significantly better during the test phase than during the practice phase). In contrast, low-Raven individuals under the other two forms of high workload (fast and load) did not improve their performance between the practice and the test phases. The right panel of Figure 4 shows that high-Raven participants in the slow condition group performed the best of all participants during both the practice and test phases.

Performance scores for the two additional tasks assigned to participants in the load condition group were quite high (>80% correct responses). An ANOVA using Raven score as the covariate showed no significant effect of Raven score on performance of the additional tasks: communications task, $F(1, 16) = 2.966$, $p = .10$; monitoring task, $F(1, 16) = 3.017$, $p = .10$. Statistical analysis revealed no significant correlations between participants' performance on the two additional tasks and their performance on the WPP task. Similarly, no significant correlation was found between their performance

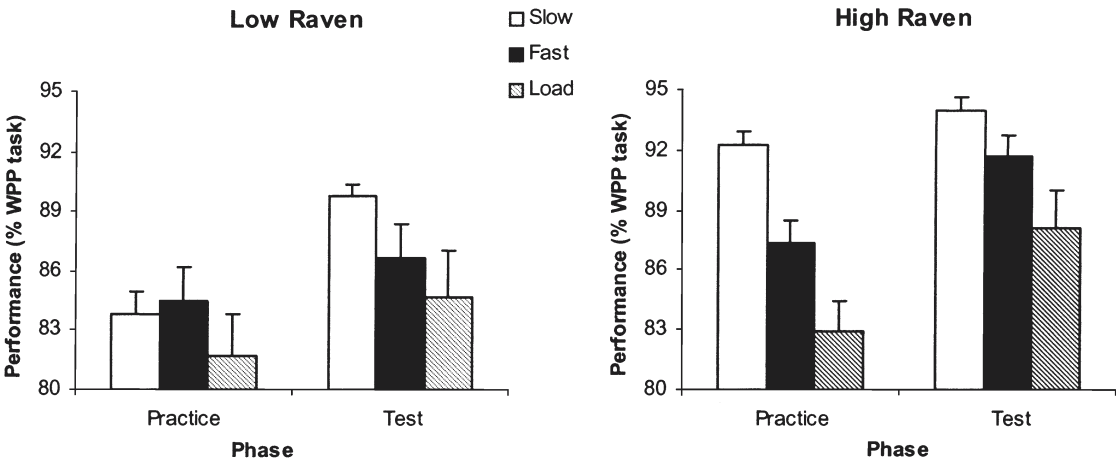

Figure 4. Average performance in the task per condition, separated into two groups (low and high Raven scores). Error bars represent 95% confidence intervals based on the pooled *MSE.*

Gonzalez,r1.qxd 3/3/05 11:04 AM Page 8

on the WPP task and the percentage of correct responses while performing each additional task, the average response time, or the between-day response time (WPP performance on Day 1 and monitoring performance: $r = -.241$, $p = .33$; WPP performance on Day 2 and monitoring performance, $r = -.053$, $p = .83$; and WPP performance on Day 3 and communications performance, $r = -.223$, $p = .37$). Moreover, differences remained insignificant even when Raven score was partialled out.

DISCUSSION AND IMPLICATIONS

The findings from this study indicate that both high task workload (in the form of time constraints and loading tasks) and low cognitive abilities (as measured by Raven score) hindered performance and transfer in DDM tasks. Moreover, these experiments demonstrate that high workload had a greater effect on individuals with low cognitive ability than on individuals with high cognitive ability.

While interpreting these results, one must remember that the study groups differed only in regard to the training phase conditions, not the testing phase conditions. Although participants ran the DDM simulation under different types of workload during the training phase, all participants completed the same number of trials (six) while under the same high workload (fast: 8 min/trial) during the testing phase.

The results suggest that low workload during training enabled participants to improve their performance more markedly after transfer to high workload than in the case of individuals who trained under high-workload conditions (either time constraints or loading tasks). Furthermore, participants in the low-workload (slow) condition completed a total of 4 trials during the training phase, whereas participants in one of the fast conditions ran the simulation 12 times during the training phase. As indicated by the higher number of trials in the fast condition and the nonsignificant difference between this and the load condition, this detrimental effect of high workload occurred independently of the number of practice trials. It is also possible, however, that a higher number of practice trials in the fast condition produced slightly better (but nonsignificant) performance as compared with the load condition. This particular hypothesis about the effect of time constraints and amount of practice has been tested in a different but related study (Gonzalez, 2004).

The study also demonstrates that the effect of task workload depended on the available cognitive resources of the decision maker. Regression analysis revealed that Raven score was a good covariate in this study and a significant predictor of performance. The analyses by condition showed a significant effect of Raven score under all conditions and an interaction between Raven score and study phase in slow condition participants only. Comparisons between the slow condition group and each of the other two high-workload groups also indicated a significant effect of Raven score and an interaction between Raven score and phase, but a comparison between the fast and load condition groups revealed no difference. These results indicate that both forms of high workload were similarly detrimental to performance improvements and were dependent on the decision makers' cognitive abilities. Only low task workload during practice helped individuals with different cognitive abilities learn to deal with high workload during testing. The analysis of individuals with different levels of cognitive abilities suggests that individuals with high cognitive abilities performed well during both the training and test phases, whereas individuals with low cognitive abilities improved their performance during the practice phase and improved even more during the test phase.

The results also support a possible link between general intelligence and learning in real-time DDM tasks (Rigas et al., 2002). This is an important result because the majority of DDM research indicates no significant associations between performance in microworlds and scores on intelligence tests, and this has led many researchers to question the validity of using fluid intelligence, also known as Gf, to predict performance in real-world pursuits (Dorner, Kreuzig, Reither, & Staudel, 1983; Omodei & Wearing, 1995; Putz-Osterloh & Luer, 1981; Rigas & Brehmer, 1999; Staudel, 1987; Strohschneider, 1986, 1991). Further, the findings suggest that context-free, simple tests such as the Raven Standard Progressive Matrices Test can serve as predictors of performance in complex DDM

Gonzalez,r1.qxd 3/3/05 11:04 AM Page 9

tasks. Successful performance on this type of test must require that test takers tap into abstract cognitive demands similar to those needed to resolve complex DDM problems (Joslyn & Hunt, 1998).

Limitations

There are several limitations of this study. First, this research addressed two forms of task workload exclusively. In general, it is expected that any kind of workload would limit the amount of time to process information and hinder information processing, especially in complex dynamic tasks. Thus the interaction with cognitive abilities found in this study should apply to other forms of task workload, although more research using other forms of workload is necessary.

Second, this study used one measure of cognitive abilities, the Raven Standard Progressive Matrices Test. It is possible that different tasks demand different cognitive abilities and that some of them are not captured by the Raven test. In terms of DDM, however, the Raven test might be an adequate measure, as research reported by Carpenter et al. (1990) suggests. The Raven test requires participants to identify how different figural attributes of a problem correspond with one another. This process is difficult because it frequently forces users to abandon hypotheses because of the ambiguous correspondence of cues. Thus the Raven test measures individuals' abilities to deal with novelty and to adapt to new cognitive problems (Carpenter et al.), an essential characteristic in DDM. Thus, although such skills seem integral to individuals' ability to perform DDM tasks well, other abilities (e.g., perceptual speed and psychomotor ability) also may influence individuals' skill acquisition in complex tasks (Ackerman, 1988), and their interactions with workload need to be explored.

Finally, this study used one DDM task. Because DDM tasks have very specific abstract characteristics that are generalizable across many domains (Gonzalez, Vanyukov, & Martin, 2005), it is expected that results from this study could be reproduced with other DDM tasks; again, however, more empirical support is needed here.

Practical Implications

The findings generated by this study could help to improve understanding of both the learning process during DDM and the design of training protocols. First, during this study decision makers' performance under high task workloads during the test phase was directly determined by the workload under which the individuals learned the task during the practice phase. Intuitively, one might think that training under conditions as close as possible to the final task conditions would be ideal – that is, that the more closely the training task mirrors the final task, the better the decision makers will perform under realistic conditions. However, the findings from this study suggest that intuition is misleading in this case. Training individuals under low-workload conditions resulted in better performance under high workload. Although some might question the validity of comparing learning and performance during the WPP simulation to the learning and performance of dynamic real-world tasks, years of research in psychology have demonstrated that the effects of real-world situations can be evaluated in controlled laboratory experiments. DDM tasks, such as the WPP simulation, often are complex and incorporate many abstract and essential characteristics of real-world situations that can be tested in the laboratory (Omodei & Wearing, 1995).

This study also suggests that it is possible to predict individuals' performance under high task workload by measuring their cognitive abilities. Researchers conducting human factors research seldom gather data on the cognitive abilities of study participants. Workload has often been considered an exogenous variable that is determined by task demands and is independent of human abilities. This study, however, showed that the effect of task demands on both performance and learning depends highly on individuals' cognitive abilities. Individuals with low cognitive abilities seem to be more sensitive to workload during learning than do individuals with high cognitive abilities. The former trainees may perform better if training begins with a cognitively less demanding version of the final tasks.

Gonzalez,r1.qxd 3/3/05 11:04 AM Page 10

ACKNOWLEDGMENTS

I would like to thank Dr. Javier Lerch for his help and support during the development of these experiments. I also would like to thank anonymous reviewers who helped improve this paper. 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 by the Office of Naval Research (N00014-01-10677).

REFERENCES

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117,* 288–318.

Ackerman, P. L., Beier, M. E., & Boyle, M. D. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General, 131,* 567–589.

Borkowski, J. G., & Burke, J. E. (1996). Theories, models, and measurements of executive functioning: An information processing perspective. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory, and executive function* (pp. 235–261). Baltimore: Paul H. Brookes.

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.

Comstock, J. R., & Arnegard, R. J. (1992). *The multi-attribute task battery for human operator workload and strategic behavior research* (NASA Tech. Memorandum 10417). Hampton, VA: NASA Langley Research Center.

Damos, D. L., & Wickens, C. D. (1980). The identification and transfer of timesharing skills. *Acta Psychologica, 46,* 15–39.

Dorner, D., Kreuzig, H. W., Reither, F., & Staudel, T. (1983). *Lohhausen: Vom Umgang mit Unbestimmtheit und Komplexitat* [Lohhausen: On dealing with uncertainty and complexity]. Bern, Switzerland: Hans Huber.

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

Gonzalez, C. (2004). Learning to make decisions in dynamic environments: Effects of time constraints and cognitive abilities. *Human Factors, 46,* 449–460.

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science, 27,* 591–635.

Gonzalez, C., Vanyukov, P., & Martin, M. (2005). The use of microworlds to study dynamic decision making. *Computers in Human Behavior, 21,* 273–286.

Gopher, D. (1993). The skill of attention control: Acquisition and execution of attentional strategies. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV* (pp. 239–322). New York: Erlbaum.

Gopher, D., & Donchin, E. (1986). Workload – An examination of the concept. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 2, pp. 1–49). New York: Wiley.

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

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. (1995). Decision making in a dynamic situation: The effect of false alarms and time pressure. *Journal of Behavioral Decision Making, 8,* 181–200.

Kerstholt, J. H., & Raaijmakers, J. G. W. (1997). Decision making in dynamic task environments. In R. Ranyard, R. W. Crozier, & O. Svenson (Eds.), *Decision making: Cognitive models and explanations* (pp. 205–217). London: Routledge.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14,* 389–433.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanism. *Psychology Review, 104,* 3–65.

Omodei, M. M., & Wearing, A. J. (1995). The Fire Chief microworld generating program: An illustration of computer-simulated microworlds as an experimental paradigm for studying complex decision-making behavior. *Behavior Research Methods, Instruments, and Computers, 27,* 303–316.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* New York: Cambridge University Press.

Putz-Osterloh, W., & Luer, G. (1981). Uber die Vorhersagbarkeit komplexer Problemloseleistungen durch Ergebnisse in einem Intelligenztest [On whether results from a test of intelligence can predict problem solving performance]. *Zeitschrift fur Experimentelle und Angewandte Psychologie, 28,* 309–334.

Raven, J. C., Court, J. H., & Raven, J. (1977). *Standard progressive matrices.* London: H. K. Lewis.

Rigas, G., & Brehmer, B. (1999). *Mental processes in intelligence tests and dynamics decision making tasks.* London: Erlbaum.

Rigas, G., Carling, E., & Brehmer, B. (2002). Reliability and validity of performance measures in microworlds. *Intelligence, 30,* 463–480.

Staudel, T. (1987). *Problemlosen, Emotionen und Kompetenz* [Problem solving, emotions and competence]. Regensburg, Germany: Roderer.

Strohschneider, S. (1986). Zur Stabilitat und Validitat von Handeln in komplexen Realitatsbereichen [On the stability and validity of complex problem solving behavior]. *Sprache und Kognition, 5,* 42–48.

Strohschneider, S. (1991). Problemloesen und Intelligenz: Ueber die Effekte der Konkretisierung komplexer Probleme [Problem solving and intelligence: The effects of problem concreteness]. *Diagnostica, 37,* 353–371.

Verwey, W., & Veltman, H. A. (1996). Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. *Journal of Experimental Psychology: Applied, 2,* 270–285.

Wickens, C. D. (1980). The structure of processing resources. In R. Nickerson & R. Pew (Eds.), *Attention and performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (1987). Attention. In P. A. Hancock (Ed.), *Human factors psychology* (pp. 29–90). New York: Elsevier Science.

Wickens, C. D. (1991). Processing resources and attention. In D. L. Damos (Ed.), *Multiple-task performance* (pp. 3–34). Bristol, PA: Taylor & Francis.

Cleotilde Gonzalez is an assistant professor in the Department of Social and Decision Sciences 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 12, 2003
Date accepted: May 13, 2004