short report

Strategy selection for cognitive skill acquisition depends on task demands and working memory capacity

Scott R. Hinze a,*, Michael F. Bunting b, James W. Pellegrino a

a Department of Psychology, University of Illinois at Chicago, M/C 285 1007 W. Harrison St., Chicago, IL 60607-7135, United States
b University of Maryland Center for Advanced Study of Language, P.O. Box 25, College Park, MD 20742, United States

A R T I C L E   I N F O
Article history:
Received 10 July 2008
Received in revised form 13 July 2009
Accepted 22 July 2009

Keywords:
Cognitive skill acquisition
Working memory capacity
Strategy use
Procedural learning

A B S T R A C T
The involvement of working memory capacity (WMC) in rule-based cognitive skill acquisition is well-established, but the duration of its involvement and its role in learning strategy selection are less certain. Participants (N = 610) learned four logic rules, their corresponding symbols, or logic gates, and the appropriate input-output combinations in three-gate circuit patterns. Participants practiced 120 repetitions of each rule (480 total gates) over the course of 10 training blocks. Memory load varied between subjects. The confluence of task demands and individual differences in WMC (N = 518) dramatically affected speed-accuracy tradeoffs and strategic use of a computerized help function. Cluster analysis revealed five distinct groups of participants based on the combination of response accuracy, latency, and help use. Some groups with moderate to high mean WMC acquired the task with predictable performance patterns. Other groups, prevalently under a memory load and with low mean WMC, failed to learn or overused help.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Cognitive abilities, task constraints and strategy use influence cognitive skill acquisition (CSA). Although these variables frequently appear in the literature, the study of their confluence is uncommon and neglected. Here we explore their interaction on learning a novel and complex logic-gate task (LGT). The LGT requires using logical functions, based on Boolean mathematics, to determine a unitary output (0 or 1) from binary input combinations (Gitomer, 1988). Four input combinations (0–0, 1–0, 0–1, 1–1) and four gate-types (AND, OR, NAND, NOR) yield 16 gate-by-input combinations. Table 1 shows the truth table of gates, rules and input–output combinations in the LG task. After an introduction to the rules, individuals practiced these gates with an option to use a “help” key to refresh their memory as needed.

Working memory (WM), a fundamental cognitive ability, is essential to CSA (e.g., Kieras, Meyer, Mueller, & Seymour, 1999; Kirschner, 2002). In Anderson’s (1982) ACT model, for example, WM holds all declarative rules and productions until propositions are compiled, strengthened, and eventually automatized through practice and chunking (Rosenbloom & Newell, 1987; Wolitz, 1988). Consistent with this assumption, individual differences in working memory capacity (WMC) and related general cognitive abilities predict skilled performance. These abilities correspond to the rate and/or accuracy of early CSA (Ackerman, 1986, 1988; Law, 2000; Wolitz, 1988), and even performance following considerable practice (Kyllonen & Stephens, 1990; Shute, 1991). One purpose of this study is to determine the extent to which complex WM span tasks predict performance during practice.

According to Engle (2002, Unsworth & Engle, 2007), WMC measures capture individual differences in attentional control, and attentional control is critical for complex CSA. We predict that the demand for attentional control is greatest during early learning when memory representations of the task goal place a high load on WM. Cognitive load manipulations can moderate these effects (Bunting, 2006). We varied cognitive load with two separate manipulations to determine whether load will have general effects on performance and/or whether participants will respond with different strategies depending on cognitive load. We expected that cognitive load would have an independent influence on performance and interact with WM, maintaining the WM–performance relationship over practice (cf. Ackerman, 1986; Wolitz, 1988).

Are load and WM effects in CSA due to more efficient general processing, different strategy choices, or both? Strategy use can profoundly influence CSA through within-task changes (Delaney, Reder, Staszewski, & Ritter, 1998), direct manipulations of between-subjects differences (Doane, Sohn & Schreiber, 1999; Doane, Alderton, Sohn & Pellegrino, 1996), and spontaneous strategy choices (Siegel, 1988). During data collection, we noticed qualitative differences in participant behavior and performance. Some participants performed well while others performed poorly, and in both cases did so with either great or little effort. Participants also used the help key to varying degrees; some even used it (almost) exclusively. Shute (1991) demonstrated that individuals often utilize help opportunities ineffectively.
either under-using help despite errors or over-using help to prevent errors, leading to poor learning outcomes. We explore similar dynamics here. We further question whether person and/or task factors mediate these strategies.

Strategies have different resource demands and effectiveness. For instance, memory retrieval is less demanding than counting in math (Siegler, 1988), and knowledge telling is less demanding than more sophisticated writing skills (McCutchen, 1988). On the LGT, different ways of responding entail different resource requirements: memory retrieval is less demanding than logical reasoning but depends upon knowledge availability in long-term memory. Using computerized help is also less demanding than logical reasoning. Varying resource demands should affect tradeoffs in accuracy, response time and help use.

Learners’ cognitive ability and external load may affect strategy availability and a person’s ability to use them. Strategy choices affect moderate-ability individuals more than extremely high or low-ability individuals (Sohn, Doane, & Garrison, 2006). Different training strategies can produce similar mean performance, but very different relationships with WM (Law, Morrin, & Pellegrino, 1995). Critically, WMC can influence strategy selection, as high WMC individuals more frequently choose resource-demanding strategies than low WMC individuals, even to the detriment of efficient task performance (Beilock & DeCaro, 2007). We expected that WMC and cognitive load should influence responses on the LGT. High WMC individuals and those under low load should be able to use resource-demanding reasoning processes, while low WMC individuals and those under high load may look for less demanding strategies like help use and memory retrieval (even when unsuccessful). One goal was to determine whether general rules define the relationship between WMC, Load, and CSA, or whether these relationships are more complex.

2. Method

2.1. Participants

Psychology subject-pool undergraduates (N = 615; 217 male), sampled over a three-year period, received nominal course credit for completing LGT and WMC tasks but due to imperfect retention, 518 have WMC data. All were native English speakers with normal or corrected-to-normal vision and unimpaired hands for speeded keyboard responses. None reported prior experience with logic gates.

Table 1

<table>
<thead>
<tr>
<th>Gate</th>
<th>Image</th>
<th>Rule</th>
<th>Input combinations (output below)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AND</td>
<td><img src="image" alt="AND gate" /></td>
<td>The output is true if both inputs are true; otherwise, the output is false.</td>
<td>0 0 0 1</td>
</tr>
<tr>
<td>OR</td>
<td><img src="image" alt="OR gate" /></td>
<td>The output is true if either or both inputs are true; otherwise, the output is false.</td>
<td>0 1 1 1</td>
</tr>
<tr>
<td>NAND</td>
<td><img src="image" alt="NAND gate" /></td>
<td>The opposite of AND. The output is false if both inputs are true; otherwise, the output is true.</td>
<td>1 1 1 0</td>
</tr>
<tr>
<td>NOR</td>
<td><img src="image" alt="NOR gate" /></td>
<td>The opposite of OR. The output is true if both of the inputs are false; otherwise, the output is false.</td>
<td>1 0 0 0</td>
</tr>
</tbody>
</table>
recall scores serves as an index of WMC (see Conway et al., 2005). Because the scores were highly correlated in our sample ($r = .65, p < .001$), they were averaged for each participant.

### 2.3. Procedure

Those in the first three semesters of data collection completed the LGT in a 60 min session and the WM tasks in a 30 min session. All others completed them in a single 120 min session with another battery. Breaks occurred at regular intervals, with no apparent differences between cycles.

### 3. Results and discussion

Response errors and latency likely reflect different aspects of CSA. Fewer errors over practice suggest the acquisition of declarative knowledge and initial proceduralization. Shorter response latency over time reflects efficient processing and proceduralization. Strategic help use, which targets specific items or rules where knowledge is lacking, can indicate successful monitoring of the contents of declarative memory and learning from feedback.

#### 3.1. Speed–accuracy-help tradeoffs

Participants can apply strategies to optimize speed, accuracy, or both. Thus, assessing learners’ speed or accuracy alone is insufficient. Here, tradeoffs in speed, accuracy, and help were possible. Because performance changed over time, strategies could shift during learning.

A taxonomy of the strategies used during CSA was derived with exploratory TwoStep cluster technique, which groups cases based on continuous or categorical attributes (Chiu, Fang, Chen, Wang, & Jeris, 2001; Everitt, Landau, & Leese, 2001). We grouped individuals by the speed, accuracy, and help use tradeoffs they used during CSA. Exploratory cluster analysis is appropriate when prior knowledge cannot inform estimates of the number of clusters, which is the case with confirmatory techniques like latent-profile analysis.

Gate 3 is the critical gate in each trial (i.e., the locus of the load effect is there), so these data were used to compute the cluster variables. Participants’ initial performance (Block 1 mean), final performance (Block 10 mean), overall mean performance, and change in performance over time (i.e., linear slope) were computed for the accuracy, reaction time and help data and used to determine cluster membership.

The cluster analysis yielded a five-cluster solution. Table 2 shows the cases per cluster and a description of each cluster’s dominant strategy tradeoffs. The cluster solution included 542 of the original 614 cases; 73 were excluded due to missing latency data on Block 1 and/or Block 10.

#### Table 2

<table>
<thead>
<tr>
<th>Cluster</th>
<th>N</th>
<th>%</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>173</td>
<td>28</td>
<td>Optimal learners</td>
<td>Accuracy started high and remained high, reaction time improved with practice, and low reliance on help.</td>
</tr>
<tr>
<td>2</td>
<td>120</td>
<td>19</td>
<td>Slow starters</td>
<td>Accuracy–speed tradeoff favoring accurate responses at the expense of long reaction times but low help use.</td>
</tr>
<tr>
<td>3</td>
<td>129</td>
<td>21</td>
<td>Help users</td>
<td>Accuracy–speed tradeoff favoring accurate responses with initially high reliance on help.</td>
</tr>
<tr>
<td>4</td>
<td>142</td>
<td>23</td>
<td>Poor learners</td>
<td>Reaction time improved with practice but accuracy started low and stayed low.</td>
</tr>
<tr>
<td>5</td>
<td>51</td>
<td>8</td>
<td>Help abusers</td>
<td>Used help (almost) exclusively throughout task.</td>
</tr>
</tbody>
</table>

Total: 615 cases

Note: Ns before the addition of excluded cases to Clusters 3, 4, and 5 were 107, 119, and 26, respectively.

### Fig. 2.

Performance characteristics of each cluster: mean accuracy (left column) for trials on which help was not used, mean latency (middle column) for trials on which responses were accurate without help use, mean help use on all trials. Help abusers’ accuracy and latency data were excluded due to limited responses without help. Error bars represent 95% confidence intervals.

### Fig. 3.

Mean standardized working memory capacity for each cluster. Error bars represent 95% confidence intervals.
(participants who used help on 100% of the trials or made 100% errors on attempted trials had missing data because latency data was computed for accurate responses only). We manually assigned these missing cases to clusters as follows: (1) 25 used help on 100% of the Block 10 trials and were added to Cluster 5, (2) 23 were added to Cluster 4 because their Block 10 mean accuracy was within ½ SD of that for Cluster 4 members, and (3) 25 were added to Cluster 3 because their Block 10 mean accuracy was within ½ SD of that for Cluster 3 members.

The strategy tradeoffs in Table 2 are based on the performance variables by cluster membership in Fig. 2. The cluster analysis yielded meaningful and interpretable groups. Clearly, no single dependent variable can adequately characterize performance. For instance, optimal learners and help users had response times with similar means and slopes, but their accuracy and help use differed. These qualitative groups match well with those identified by Shute (1991), including both effective and counter-productive allocations of resources. In the next section, we test our prediction that cognitive abilities and cognitive load influence these strategy choices.

### 3.2. Predicting group membership

Nominal logistic regression analysis allows a test of the relationship between discrete or continuous independent variables on discrete dependent variables (Tabachnick & Fidell, 2001). There was one discrete dependent variable (cluster membership) with five levels, and three between-subject independent variables: WMC (continuous and standardized), memory for output [binary (load and no-load)], and learning set [binary (partial-set and full-set)]. Participants for whom WMC data were unavailable were excluded from all logistic regression analyses (final N = 518).  

A model using the main effects of WMC, memory for output, and learning set was statistically reliable according to a likelihood ratio test, \( \chi^2(12) = 54.94, p < .001 \). While a full factorial model using interaction terms was also reliable, the interaction terms were not significant, nor was the difference between the models (\( \chi^2(4) = 5.80, p = .21 \)), so the main-effects model is preferred. There were main effects for WMC and memory for output and a trend for learning set, \( \chi^2(4) = 35.74, p = .001 \), \( \chi^2(4) = 11.24, p = .02 \), and \( \chi^2(4) = 7.56, p = .11 \), respectively.

By visual inspection of Fig. 3, optimal learners and slow starters had better-than-average WMC, the help users had average WMC, and the poor learners and help abusers had lower-than-average WMC. Wald tests, which compare the odds of being in one cluster versus another based on the predictor variable, confirmed these observations. Higher WMC individuals were more likely to be optimal learners than help users (Wald = 11.30, \( \beta = .63 \), CI = .48:.83), poor learners (Wald = 16.37, \( \beta = .58 \), CI = .45:.76), or help abusers (Wald = 18.82, \( \beta = .44 \), CI = .31:64). WMC did not differentiate slow starters from optimal learners (Wald = .01), but did differentiate slow starters from help users (Wald = 8.49, \( \beta = .84 \), CI = .48:.87), poor learners (Wald = 12.36, \( \beta = .59 \), CI = .44:.79), and help abusers (Wald = 16.10, \( \beta = .45 \), CI = .30:66). Help users did not differ from poor learners, but they were higher in WMC than help abusers (Wald = 3.74, \( \beta = .70 \), CI = .49:1.01). WMC did not differentiate poor learners from help abusers.

Fig. 4 shows the percentage of participants in the load condition for each cluster. Most help abusers were in the load condition (70.6%). Thus, memory for output differentiated help abusers from optimal learners (Wald = 6.86, \( \beta = 2.78 \), CI = 1.29:5.98), slow starters (Wald = 4.78, \( \beta = 2.45 \), CI = 1.10:5.47), help users (Wald = 9.86, \( \beta = 3.47 \), CI = 1.60:7.53), and poor learners (Wald = 4.49, \( \beta = 2.27 \), CI = 1.06:4.84). Load failed to differentiate any of the other groups, though optimal learners (43.9%) and help users (45.7%) had the lowest percentage of participants in the load group.

Fig. 4 also shows the percentage of participants in the full-set condition by cluster. Wald tests suggest this marginal effect is interpretable. The majority of the help abusers (66.7%) were in the full-set condition, while the majority of the help users were in the partial-set condition (54.3%). Thus, the learning set condition differentiated help users from help abusers (Wald = 5.08, \( \beta = 2.35 \), CI = 1.12:4.93). Learning set did not differentiate the remaining groups.

These results effectively show that participants’ choice of different resource-allocation strategies is not random. Rather, these differences can be predicted, in part, by both WMC and the cognitive load built into the task.

### 3.3. Individual differences

Research (e.g. Ackerman, 1986, 1988; Woltz, 1988) suggests that WMC’s correlation with accuracy, though initially high, should attenuate over practice. We also predicted that these correlations would be greater under high cognitive load conditions. However, the cluster analysis and logistic regression data suggest the relationship may be more complex.

---

1. Although a smaller sample is used for the remaining analyses, cluster analysis performed with the smaller sample yielded a similar 5-factor solution, so the full model is presented here for a more complete descriptive presentation.
These strategic differences were interpretable through cluster analysis, and reflected strategic resource-allocation tradeoffs. Optimal learners acquired the logic rules quickly without help and were able to use practice to improve response latency. Other participants (slow starters and help users) used additional processing time or help to reinforce learning. Still others (poor learners and help abusers) did not demonstrate effective acquisition and proceduralization of the logic rules. Multiple factors could account for these differences in resource-allocation. For instance, some participants may not have fully understood the training on logical rules or the nature of the task. And, it is tempting to write-off poor learners and help abusers as unmotivated or as cheaters. However, WMC and experimenter-imposed memory load partially predicted these strategic differences. These results have practical and theoretical implications.

Practically, these results suggest that task demands and WMC may influence not only efficient performance, but also more general learning strategies. Individuals reaching cognitive limits may reduce cognitive demands by using help, or by responding quickly without much effort, each at the expense of learning. Reducing cognitive load may allow these participants to engage in more effective learning strategies. Alternatively, individuals with available resources who are initially unsuccessful may choose to “plug away” at the task, at the expense of training time, rather than use help strategically, as seen with the slow starters.

Theories make different predictions for the duration and relative importance of WM over the course of acquisition and proceduralization. Our data are consistent with research and theory-based predictions that WM is initially critical to the acquisition process, but the role of WM wanes given efficient proceduralized responses. When participants reached high levels of performance and proceduralization occurred, for example in Clusters 1–3, individual differences in WMC became less important over practice. However, when participants favored speed over accuracy (poor learners) or were allowed to “game” the system (help abusers), the overall pattern changed. Low WMC individuals were less likely to reach later stages of skill, so the WMC–performance relationship remained. In addition, task demands such as the imposition of a memory load and the task set manipulations prolonged the dependence of performance on WMC, even among those who were successful. The influence of task and strategic factors suggest it may not be surprising that some authors have found a sustained influence of cognitive abilities after much practice (e.g. Kyllonen & Stephens, 1990; Shute, 1991), while others have found a reduction in influence with proceduralization (e.g. Ackerman, 1986; Woltz, 1988). These data are also consistent with the emerging literature suggesting that WMC may influence processing strategies in addition to processing efficiency (e.g. Beilock & DeCaro, 2007).

A limitation of this research is that participants had only 1 h of practice. In contrast, Carlson, Sullivan and Schneider (1989) had participants engage in 20 h of logic-gate practice. It would be interesting to know the effect of extensive practice on strategy use. We would expect a convergence of the participants in the optimal learners, slow starters, and help users groups. Although these groups had obvious differences in help use and the nature of their speed-accuracy tradeoffs, they demonstrated effective learning and should all be expected to reach asymptotic performance. We have no reason to anticipate an improvement in performance for the poor learners and help abusers under challenging task demands.


