

## Strategy shifts in classification skill acquisition: Does memory retrieval dominate rule use?

LYLE E. BOURNE, JR., ALICE F. HEALY, JAMES A. KOLE, and SUSAN M. GRAHAM  
*University of Colorado, Boulder, Colorado*

In two experiments, we demonstrated two types of strategies (rule-based and memory-based) and strategy transitions within the same binary classification task. The strategy that dominated later in practice depended on the difficulty of the operative classification rule and on the salience of the cue for that rule. Accuracy increased over practice trials, and response times were faster for the dominant strategy, either rule or memory. Rule retention was superior to stimulus item retention, so that, even for participants who preferred a memory-based strategy, a rule-based strategy dominated at least temporarily after a 1-week interval. Strategy use over trials and the retention interval reflected a given task's affordance of a shift between rule- and memory-based processes.

Strategies adopted by participants during skill acquisition do not necessarily remain stable over training or practice trials, but rather often shift in the direction of speeding up responses and easing the cognitive load (Schunn & Reder, 2001). A number of formal models, notably by Logan (1988) and Rickard (1997), have been proposed to account for strategy shifts. These models have focused on a shift from performance based on a rule (algorithm, calculation) to performance based on the retrieval of previously experienced items from memory. In these models, retrieval from long-term memory is viewed as a single-step, possibly automatic solution procedure for problems that might otherwise require multiple rule-based computations. In tasks to which these models have been applied, participants are first instructed in the rule or algorithm that governs responses. Thus, although there are certain fundamental differences between them, both the Logan and Rickard models assume that performance is under control of an algorithm-based strategy at the outset and that, at some point in practice, there is a transition in control from a strategy based on the algorithm to a strategy based on instance-memory retrieval.

The Logan (1988) and Rickard (1997) models may be limited by the tasks to which they have been applied (e.g., alphabet arithmetic). In these tasks, only two strategies are possible. Participants are given the rule (or algorithm)

in advance and told to use the rule-based strategy at the outset. Under these conditions, guessing is unlikely and a transition from memory retrieval to rule use is virtually impossible. In a binary classification task, using a simple self-report procedure to identify strategies, Bourne, Healy, Parker, and Rickard (1999) found that strategy shifts are not always in the rule-to-memory direction. When participants were not preinstructed in the rule or algorithm for a task, most of them reported guessing on early trials. After some significant number of repetitions of individual stimuli, most participants adopted a preferred strategy, which might or might not be memory based. Specifically, most participants preferred a rule-based strategy in what was called a *natural task condition*. This task was called natural because it involved a linguistic rule for pronunciation of English words (i.e., participants had to choose one of two pronunciations, "thee" or "thuh," for the definite article *the*, depending on a given noun or adjective; see Raymond, Fisher, & Healy, 2002, for evidence that college students are not fully knowledgeable about this rule despite its presence in all English dictionaries). In contrast, most participants eventually preferred a memory-based strategy in an *artificial task condition* (in which participants had to base their response on whether meaningless letter strings conformed to an alphabetical sequence). Overall measures (accuracy and response time [RT]) showed that performance was better under the control of the preferred strategy—either rule or memory—than under the nonpreferred strategy.

The findings by Bourne et al. (1999) suggest that extant models probably do not satisfy all tasks in which strategy use and strategy shifts occur. It may be that rule use will persist indefinitely if the rule can be applied more quickly than memory retrieval in the later stages of practice. Touron and Hertzog (2004) have interpreted the persistence of a rule strategy in the Bourne et al. experiments to suggest that strategy shifts might occur only when the processing cost of a given strategy (e.g., the difficulty of retrieving item information from memory) is offset by its

---

This research was supported in part by Army Research Institute Contracts DASW01-99-K-0002 and DASW01-03-K-0002 and Army Research Office Grant DAAG55-98-1-0214 to the University of Colorado. The fourth author also received support from the Undergraduate Research Opportunities Program at the University of Colorado. The authors wish to thank James Parker for help with the preparation of the experimental materials and the statistical analysis of Experiment 1. Correspondence should be addressed to L. E. Bourne, Jr., Department of Psychology, University of Colorado, Boulder, CO 80309-0345 (e-mail: lbourne@psych.colorado.edu).

*Note*—This article was accepted by the previous editorial team, when Colin M. MacLeod was Editor.

benefits (e.g., reduced processing demands or increased processing efficiency). They refer to the relative costs and benefits of a strategy shift as *shift affordance*. Thus, experimental manipulations that increase the relative benefit of a memory strategy over a rule-based strategy are said to increase the task's affordance of a strategy shift. The Bourne et al. data suggest that one such manipulation might be rule difficulty. The more difficult the rule is to apply, the greater is the task's affordance of a shift from rule use to memory.

Like Logan (1988) and Rickard (1997), Touron and Hertzog (2004) considered strategy shifts only in the rule-to-memory direction. However, their notion of shift affordance clearly would apply to experimental manipulations, if there are any, that change the cost/benefit ratio in favor of a rule strategy, affording a memory-to-rule strategy shift. We identified one such variable, an interpolated retention interval, in Experiment 1.

## EXPERIMENT 1

Preference for the rule strategy in the natural task condition persisted over 30 blocks of 12 training trials in the study reported by Bourne et al. (1999). That amount of practice (or exposure to repeated stimuli) might be insufficient to allow a dominant memory-based strategy to emerge (see Klapp, Boches, Trabert, & Logan, 1991). That is, the models of Logan (1988) and Rickard (1997) might not have been given a fair test in previous experiments. So, in Experiment 1, we added a second 30-block session, administered 1-week after the first session. The main purpose was to determine whether, with additional exposure to the same training stimuli, rule use would continue to dominate performance in the natural task or whether, even in this natural task, participants would eventually shift from rule use to memory retrieval.

Inclusion of a second session 1-week after the first also allowed us to determine whether the rules and instances acquired during initial training were remembered over a delay interval and used to the same degree after an opportunity for forgetting. There is evidence to suggest that rules and instances often suffer different amounts of forgetting. In a pseudo-arithmetic paradigm, Rickard (1994; see also Healy et al., 1995, and Rickard & Bourne, 1995) showed a slowing of RTs over a 6-week retention interval that could be explained entirely in terms of a switch by some participants from a rapid-acting memory strategy to a more deliberate rule strategy, a strategy shift that was attributed to the greater forgetting of instances than of the rule. Similar results were reported by Logan and Klapp (1991). From these findings, we might expect that participants, regardless of condition in Experiment 1, would evidence more rule use at the outset of Session 2 than they did at the end of Session 1. Thus, the opportunity for forgetting across the retention interval might afford a shift back from memory retrieval to rule use.

Finally, if rule difficulty accounts for the difference in strategy use and other aspects of performance in the ex-

periments of Bourne et al. (1999), then making the artificial rule simpler, and thus lessening its affordance of a rule-to-memory strategy shift, should cause performance in the artificial task to resemble performance in the natural task. In Experiment 1, therefore, we compared the artificial and natural tasks used previously to a new artificial task based on a very simple alphabetic rule, in an effort to provide a direct test of the rule difficulty hypothesis.

## Method

**Overview and Design.** There were three experimental conditions: artificial hard, artificial easy, and natural. In each condition, the participants were trained on a single set of 12 different letter strings or words presented one at a time repeatedly over 30 blocks of practice. A block is defined as one presentation of each of the 12 stimuli in a random order. The participants were asked to give a categorical response ("code"/"noncode" for the artificial conditions or "thee"/"thuh" for the natural condition) for each stimulus by pressing an appropriately labeled response key, and each response was followed immediately by ("correct"/"incorrect") feedback. The participants were not instructed in the rule that determined correct responses. For six of the stimuli, the correct response was "code" (or "thee"), and for the remaining six stimuli, the correct response was "noncode" (or "thuh"). Without any cue, every second and third block of practice was followed by a novel stimulus, for a total of 20 novel trials. Of the 20 novel stimuli that were used, 10 were "code" (or "thee") and 10 were "noncode" (or "thuh"). Responses to these novel stimuli were also followed immediately by feedback. The participant strategies were probed after the feedback was presented on every trial to determine whether the participants (1) guessed, (2) used a rule, (3) remembered the answer to the particular instance from a previous trial, or (4) did something else (other, unspecified). After the strategy response, there was a 250-msec pause, and then the next stimulus appeared. Each participant was trained and tested 1 week later on the same set of stimuli with the same procedure. The dependent measures were the RTs, the proportion of correct responses, and the percentage of strategies of each type reported for training and novel stimuli.

**Participants.** Thirty-six undergraduate University of Colorado students participated. Twenty-five volunteered for credit in an introductory psychology course, and 11 were paid \$12 for their service. All participants were native English speakers. The participants were assigned by a fixed rotation to the three conditions, with 12 in each condition.

**Materials and Apparatus.** Five Macintosh Quadra 605 computers were used in conjunction with a PsyScope program for presenting stimuli and collecting data. The participants used specially marked keys to indicate whether the stimuli were "code"/"noncode" or "thee"/"thuh" and to select one of four strategy options.

Three-letter strings were the stimuli in the artificial hard condition. By the rule used in that condition, a letter string was correctly categorized as code only if it could be rearranged to correspond to a sequence of adjacent letters in the alphabet (e.g., RQP is code because it can be rearranged as PQR, whereas FGJ is noncode). Stimuli were constructed according to the procedures described in Bourne et al. (1999). Two-letter strings were the stimuli in the artificial easy condition. In that condition, all strings were presented in alphabetical order. The code strings were successive letters in the alphabet (e.g., DE), whereas the noncode strings were two letters that were separated by at least 10 positions in the alphabet (e.g., GR). Thus, the rule in this condition is the simplest possible version of the alphabetical rule used in the artificial hard condition. The same high-frequency words (nouns and adjectives) used by Bourne et al. were the stimuli in the natural condition. The participants were asked to indicate whether they would pronounce the definite article *the* as "thee" or "thuh" preceding the stimulus. The correct keypress

was “thee” for the stimulus words that began with a vowel and was “thuh” for the stimulus words that began with a consonant.

**Procedure.** The participants were tested in individual quiet rooms, free from outside distractions. The instructions, which outlined the procedure in detail, were presented to the participants on the computer monitor. Those instructions included the following main points: (1) The stimuli were strings of three letters (or strings of two letters or words) presented one at a time on a computer screen. (2) The same 12 stimuli were presented repeatedly over 30 blocks, with each block including one exposure to each of the 12 stimuli. (3) The task was to indicate which response, “code” or “noncode,” was correct for each letter string or which response, “thee” or “thuh,” went with each word. (4) The participants were told that a relatively simple rule would discriminate between the two categories of stimuli. (5) Each category response was followed by correct/incorrect feedback. (6) After each category response and feedback, the participant was asked to indicate which of four strategies was used on that trial. Specifically, they were told, “You should select the ‘Guess’ option if you made your response having no idea whether the preceding stimulus really belonged to one or the other category. Select the ‘Rule’ option if you discovered a useful rule and made your response on the basis of it. Select ‘Memory’ if you remembered the category for the preceding stimulus from an earlier experience with it IN THIS EXPERIMENT. Select ‘Other’ if none of the first three options applies.” (7) The participants were told that, on occasion, a new stimulus would appear that they had not seen before, but that the category task was the same for those stimuli. The instructions were the same in all conditions except for appropriate adjustments in the description of stimuli and responses.

## Results

**Accuracy.** In order to relate strategy use to performance, it is important to show that response accuracy improves systematically over trials. Indeed, the proportion of correct responses for each of the three conditions within both sessions increased across blocks of training trials [ $F(29,957) = 11.12$ ,  $MS_e = 0.0072$ ,  $p < .001$ ; see Figure 1]. Consistent with the expected forgetting over a retention interval, there was a slight loss of accuracy on

the first few blocks of Session 2, reflected in a significant main effect of session [ $F(1,33) = 12.50$ ,  $MS_e = 0.0588$ ,  $p = .002$ ] and a significant interaction of block  $\times$  session [ $F(29,957) = 5.20$ ,  $MS_e = 0.0059$ ,  $p < .001$ ].

**Strategy use.** The proportion of “guesses” started relatively high, but it quickly dropped to near 0 over blocks in all three conditions (see Figure 2). The proportion of “other” strategies was low and essentially constant over blocks in both sessions. These results replicate the observations of Bourne et al. (1999).

For memory-based and rule-based strategy reports, the data differed among conditions, consistent with the rule difficulty hypothesis. In the artificial hard task, the participants showed an initial increment in rule use, which gradually gave way to the memory strategy. These are the results that would be expected if rule- and memory-based strategies compete for control of category responses, if rule use is difficult relative to memory retrieval, and if instance strength or number grows in memory with repetitions. That is, these results are consistent with the idea that this task provides rule-to-memory shift affordance (Touren & Hertzog, 2004). In the first block of the second session, however, the participants reverted to rule use as the preferred strategy and proportion memory use dropped below .2. This finding suggests that the participants forgot specific instances and/or their category assignments, but not the rule they acquired in Session 1. Thus, the passage of time affords the opportunity for a reverse strategy shift (i.e., from memory to rule), an affordance not anticipated by Touren and Hertzog. During the subsequent blocks, memory use recovered, reaching the same asymptote as in the first session, whereas rule use decreased to its Session 1 asymptote.

Rule use was relatively high at the outset in the natural task, gradually increased across blocks of trials in Ses-

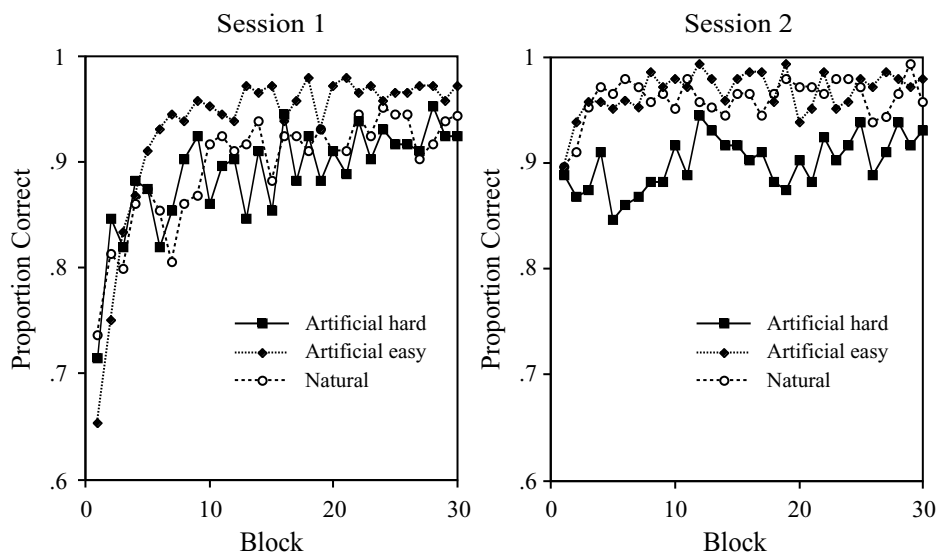


Figure 1. Proportion of correct responses as a function of condition, block, and session for the training stimuli in Experiment 1.

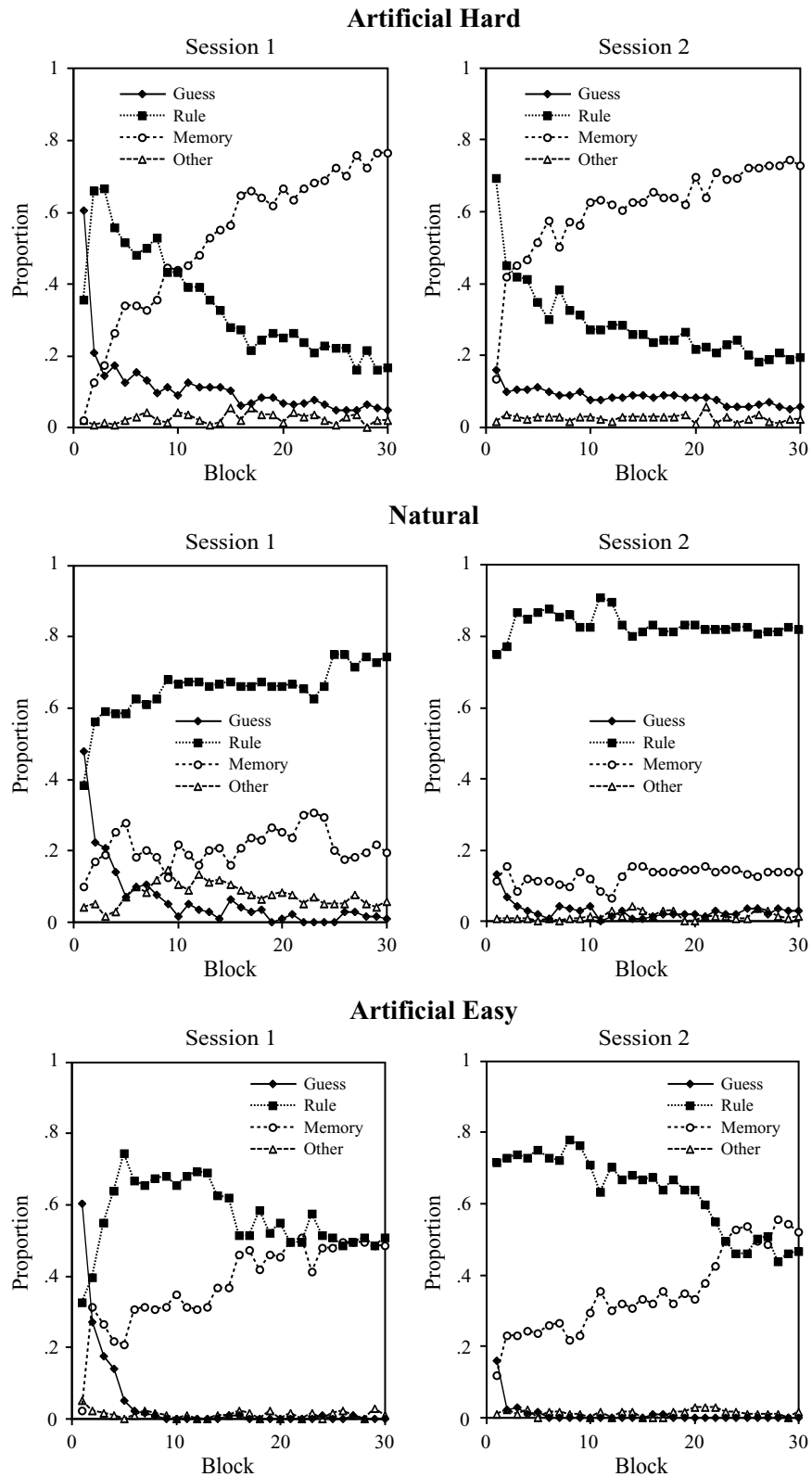


Figure 2. Proportion of rule, memory, guess, and other strategy use in both sessions as a function of block for the artificial hard, natural, and artificial easy tasks for the training stimuli in Experiment 1.

sion 1, and continued to be the dominant strategy throughout Session 2. The results for the natural and the artificial hard tasks in the first session essentially replicated previous findings (Bourne et al., 1999). These results confirm the rule difficulty hypothesis, which implies greater affordance of a shift from rule to memory with more difficult rules. Given the same pattern of results in Sessions 1 and 2, a possible transition to the memory retrieval strategy with extensive enough practice in the natural task, which is implied by the Logan (1988) and Rickard (1994) models, appears to be unlikely.

The most interesting task for testing the rule difficulty hypothesis was artificial easy, which was new to the present study. Here, the pattern of strategy use was a compromise between the patterns observed in the other two conditions. In Session 1, both rule use and memory use increased over early blocks of trials, with rule use initially dominating. Rule use then gradually declined to converge with memory use at about Block 20. Again, there was forgetting of instances but no forgetting of the rule across the 1-week retention interval separating the two sessions. In Session 2, rule use declined and memory use increased, to converge again at about Block 20.

We conducted an ANOVA to compare the three task conditions with respect to their use of the two most interesting strategies—rule and instance. The two-way interaction of strategy  $\times$  condition was significant [ $F(2,33) = 6.05$ ,  $MS_e = 10.3061$ ,  $p = .006$ ], indicating that overall rule use was preferred in the artificial easy and, especially, in the natural conditions, whereas memory use was preferred in the artificial hard condition. More important was the significant three-way interaction of condition  $\times$  block  $\times$  strategy [ $F(58,957) = 3.09$ ,  $MS_e = 0.0968$ ,  $p < .001$ ], supporting the claim that the three conditions each showed a unique pattern of strategy use as a function of block.

*Accuracy and RT as a function of strategy.* If the strategy probe procedure is valid, then the proportion of correct responses should be near chance when the guessing strategy was reported but near perfect when the other strategies were reported. Indeed, the correct response proportion was much lower (.488) for the guessing strategy than for the rule (.973) and memory (.945) strategies [ $F(2,66) = 159.50$ ,  $MS_e = 0.033$ ,  $p < .001$ ].

If the rule difficulty hypothesis is correct, there should be an interaction between condition and preferred strategy in RT. In fact, RTs were more rapid when the participants reported using the memory strategy than when they reported using the rule strategy for the artificial hard task (memory, 3.029 sec; rule, 3.163 sec), and they showed the opposite pattern for the natural task (memory, 2.992 sec; rule, 2.762 sec). The pattern for the new artificial easy task was again a compromise between those for the other two tasks, with RTs nearly equal for rule and memory trials, but with slightly faster responses on rule trials (memory, 3.002 sec; rule, 2.911 sec). Hence, in both sessions, RTs were clearly related to the dominant strategy in each task, and RTs were nearly equivalent for the two strategies where there was no consistent preference. An ANOVA using blocks as a random effect provided statistical sup-

port for the observed patterns in terms of a significant two-way interaction of condition  $\times$  strategy [ $F(2,87) = 541.98$ ,  $MS_e = 0.002$ ,  $p < .001$ ].

*Analysis of retention.* To assess the change in strategy use across the 1-week retention interval, we conducted an analysis to compare the last five blocks of Session 1 and the first five blocks of Session 2. This analysis yielded a number of significant effects, including an interpretable four-way interaction of condition, strategy type, session, and block (see Table 1) [ $F(8,132) = 3.74$ ,  $MS_e = 0.021$ ,  $p < .001$ ]. Specifically, the participants reported a higher proportion of rule use (and a corresponding lower proportion of memory use) during the first five blocks of Session 2 than during the last five blocks of Session 1 in all three conditions. However, the change from Session 1 to Session 2 was greatest in the artificial hard condition, where the participants preferred the memory strategy at the end of Session 1, and was smallest by far in the natural condition, where the participants showed a relatively low level of memory strategy use at the end of Session 1. Thus, the participants in the artificial hard condition did acquire a rule in Session 1 and were able to remember it in Session 2. However, the rule was quickly replaced as the participants reacquired the instances. These observations are consistent with the intuitively reasonable idea that participants more readily forget instances than rules across a retention interval but return to their preferred memory strategy with additional exposure to the items (see Logan & Klapp, 1991; Rickard, 1994). The frequency of rule use in the early retention trials of all task conditions supports the argument that neither rule nor memory strategy is disengaged while the other is dominant in Session 1.

**Responses to novel stimuli.** The validity of the strategy probe procedure is supported by the fact that, in all three conditions, the participants were most likely to report using a rule on novel items, followed by guessing; memory and other strategy reports were relatively infrequent (see Table 2). Because these were previously unseen items, a memory strategy report would have been inappropriate unless the novel item was highly similar to a training item and thus confused with it (see, e.g., Allen & Brooks, 1991; Johansen & Palmeri, 2002). To confirm these conclusions, we conducted an analysis of the pro-

**Table 1**  
Proportion of Rule and Memory Strategy Use on the Last Five Blocks of Session 1 and the First Five Blocks of Session 2 As a Function of Condition in Experiment 1

Condition	Session 1 Trial Block					Session 2 Trial Block				
	26	27	28	29	30	1	2	3	4	5
Artificial Hard										
Rule	.22	.16	.22	.16	.17	.69	.45	.42	.41	.35
Memory	.70	.76	.72	.76	.76	.13	.42	.45	.47	.51
Artificial Easy										
Rule	.49	.49	.51	.49	.51	.72	.73	.74	.73	.75
Memory	.49	.49	.49	.49	.49	.12	.23	.23	.24	.24
Natural										
Rule	.75	.72	.74	.73	.74	.75	.77	.87	.85	.87
Memory	.17	.18	.19	.22	.19	.11	.15	.08	.12	.11



**Table 2**  
**Proportion of Strategy Use on Training and Novel Items As a**  
**Function of Condition in Experiment 1**

Condition	Training			Novel		
	Guess	Rule	Memory	Guess	Rule	Memory
Artificial hard	.10	.31	.57	.18	.68	.10
Artificial easy	.03	.60	.37	.18	.74	.06
Natural	.05	.74	.17	.12	.79	.04

portion of times the guess, memory, and rule strategies were reported as a function of condition, session, and stimulus type (i.e., training or novel items). As shown in Table 2, and supported by the three-way interaction of these variables [ $F(4,66) = 4.05$ ,  $MS_e = 0.086$ ,  $p = .005$ ], the proportion of rule use was consistently high for the novel items in all three conditions, even though it was relatively low in frequency for the training items in the artificial hard condition.

### Discussion

At the end of both sessions, for nearly all participants in this experiment, one strategy, either memory retrieval or rule based, was dominant, and RTs were faster under the dominant strategy. The dominant strategy was rule based in the natural task and was memory based in the artificial hard task. There was no evidence of reduced rule use or a shift to memory-based performance even with extended practice in the natural task, contrary to expectations based on models of Rickard (1997) and Logan (1988). The participants were evenly divided in their preference for rule and memory strategies at the end of practice in the artificial easy task. As in the other conditions, responses were faster on the preferred strategy than on the less preferred strategy, and RTs were nearly equivalent for the two strategies where there was no consistent preference. Thus, when an artificial rule was made simpler, lessening its affordance of a rule-to-memory strategy shift (Touron & Hertzog, 2004), performance in the artificial task resembled performance in the natural task condition. This outcome would appear to offer additional support for the rule difficulty hypothesis regarding strategy use and strategy shifts. The full pattern of results from this experiment implies that rule difficulty, or ease of rule use, is the major factor determining strategy preference, strategy shifts, response accuracy, RT, and shift affordance in binary classification tasks.

In all conditions of this experiment, accuracy decreased only slightly from the end of practice in Session 1 to the beginning of practice in Session 2. However, there were substantial changes in strategy reports. In all three conditions, the participants showed an increase in rule use at the beginning of Session 2, relative to at the end of Session 1, even when the memory strategy was preferred in Session 1. These changes from memory to rule use across the retention interval reflect the greater memorability of a single rule relative to a collection of instances. They also suggest that rule strength persisted even when the memory strategy dominated performance in the second half of Session 1. These

results are similar to those obtained by Rickard (1994), who found that a large number of participants switched from a memory strategy to a rule strategy in a pseudo-arithmetic task over a 6-week retention interval, and they are consistent with Logan's (1988) assumption that both rule-based and memory retrieval strategies can be active, even when only one controls overt category responses.

### EXPERIMENT 2

When the stimuli and the rule are simple and familiar, as in the natural task condition of Experiment 1 and earlier studies (e.g., Bourne et al., 1999), rule use persists, and there is no shift from a rule-based strategy to a memory-based strategy with extended practice. We had expected that the pattern of results in the artificial easy task in Experiment 1 would reproduce that obtained in the natural task. However, accuracy, RTs, and strategy use in the artificial easy task were intermediate between those of the artificial hard and natural tasks. The rule used to define the artificial easy task clearly was not comparable in its simplicity to the rule defining the natural task, possibly because this artificial alphabetic rule required consideration of both letters of each stimulus, whereas only the initial letter determined category membership under the rule for the natural task.

Rule difficulty does seem to play a role in these results, but there are a number of other differences between the natural and artificial conditions that might also be important. For example, the stimuli for the natural task were meaningful words, and the stimuli for the artificial tasks were arbitrary two- or three-letter strings. Furthermore, the category response buttons for the natural task were labeled "thee" and "thuh," and the buttons for the artificial tasks were labeled "code" and "noncode." In Experiment 2, we explored some of these differences by using a different version of the artificial tasks. Specifically, we created an artificial version of the natural task by scrambling the letters in each stimulus word except for the first letter, which was either a vowel or a consonant. For example, the word *office* in the natural task became *ocifef* in the artificial task. If the strategy adopted by the participants focuses solely on the first letter of the stimulus, then the rules for categorizing stimuli in these tasks should be equivalent and thus equally difficult.

It is possible that the salience of the cue (or letter) to which a rule applies might also play an important role in determining strategy use. Lassaline and Logan (1993) demonstrated that cue salience varies with spatial position in a visual array. To examine this possibility, we created a second artificial task using the same scrambled words, but with letters in reverse order, such that the same vowel-consonant distinction applied but to the last letter rather than the first. In this case, the artificial equivalent of the word *office* (i.e., *ocifef*) became *fefico*. Following Lassaline and Logan, we argue that, because it requires the participants to scan to the end of the stimulus and to ignore the more prominent initial letter, the one-letter vowel-consonant rule should be more difficult to discover or

slower to use in this last-letter condition than in the first-letter condition because its cue is less salient.

## Method

**Design.** There were three experimental conditions: natural, first-letter artificial, and last-letter artificial. As in Experiment 1, in each condition, after instructions, the participants were trained on a single set of 12 letter strings or words presented on a monitor one at a time repeatedly over 30 blocks of practice. The participants were asked to give a categorical response by keypress for each stimulus (“code”/“noncode” in all conditions including the natural), and each response was followed immediately by (“correct”/“incorrect”) feedback on the monitor. The correct response was determined by the vowel/consonant rule unknown to the participant at the onset of the experiment. Every second and third block of practice was followed by a novel stimulus, for a total of 20 novel trials. Of the 20 novel stimuli that were used, 10 were code and 10 were noncode. Responses to these novel stimuli were also followed immediately by feedback. The participant strategies were probed in the same way as in Experiment 1. After the strategy response, there was a 250-msec pause, and then the next stimulus appeared. Each participant served in a single training session (no-retention session).

**Participants.** Thirty-six undergraduate University of Colorado students participated, all of whom volunteered for credit in an introductory psychology course and were native English speakers. The participants were assigned by a fixed rotation to the three conditions with 12 in each condition.

**Apparatus and Materials.** The same apparatus and computer programs were used as in Experiment 1.

The same high-frequency words (nouns and adjectives) used in the natural condition of Experiment 1 were employed in the natural condition of Experiment 2. The stimuli in the first-letter artificial condition were scrambled-letter versions of the words used in the natural condition, maintaining the same initial letter. The stimuli in the last-letter artificial condition were backward versions of the stimuli in the artificial condition.

**Procedure.** The same procedure was used as in Experiment 1. The instructions were the same as in the artificial conditions of Ex-

periment 1 and were the same for the participants in all three conditions, with the minor exception that the stimuli were referred to as *letter strings* in the artificial conditions but as *words* in the natural condition. No mention was made of the definite article (*the*) or its pronunciation in the natural condition. In addition, the participants were told that the category and strategy responses were equally important. Again, the instructions were presented to the participants on the computer monitor.

## Results

**Accuracy.** Figure 3 shows that the proportion of correct responses for the training stimuli in each of the three conditions increased systematically across blocks of trials [ $F(29,957) = 34.74$ ,  $MS_e = 0.014$ ,  $p < .001$ ]. There were no significant effects involving condition, although numerically accuracy was on average across trials higher in the natural task (.857) than in the first-letter artificial (.801) or last-letter artificial (.804) tasks.

**Strategy use.** Figure 4 summarizes the proportion of rule and memory strategy use reports by participants as a function of block for all three tasks. For the natural task, the pattern was similar to that observed in earlier experiments. How this pattern compares with those obtained in the artificial tasks is revealed by the way the performance functions cluster during approximately the last one third of trials. The upper cluster (natural rule, first-letter artificial rule, last-letter artificial memory) shows that rule use became the dominant strategy at the end of training for the natural and first-letter artificial tasks. Memory use was dominant in the last-letter artificial task. Also, the bottom cluster (natural memory, first-letter artificial memory, last-letter artificial rule) shows that rule use in the last-letter task and memory in the natural and first-letter tasks all declined with training. This outcome is

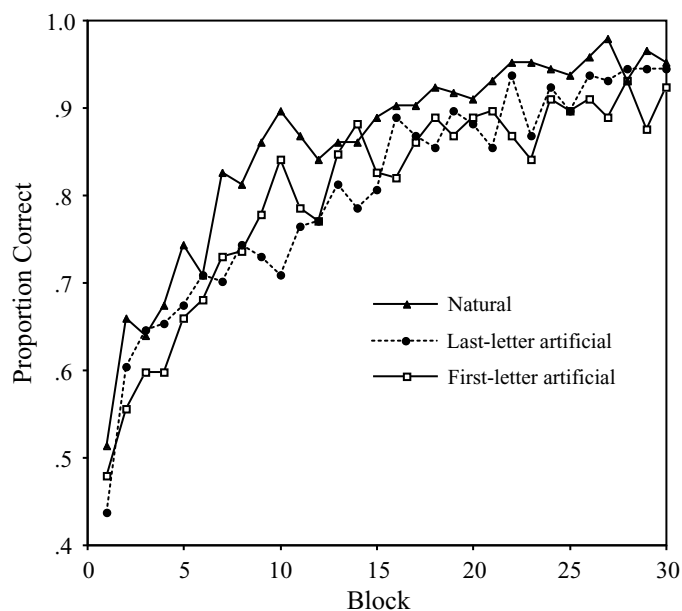
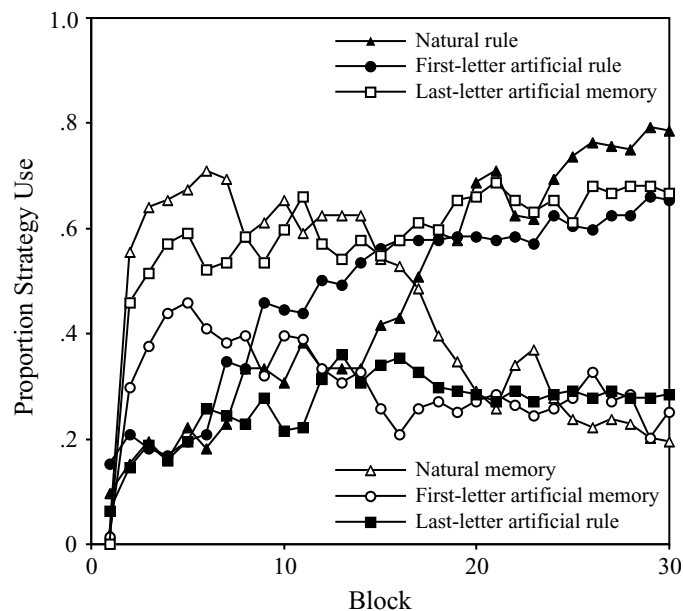


Figure 3. Proportion of correct responses as a function of condition and block for the training stimuli in Experiment 2.



**Figure 4. Proportion of rule and memory strategy use as a function of block for the natural, first-letter artificial, and last-letter artificial tasks for the training stimuli in Experiment 2.**

consistent with our hypothesis that the dominant strategy depends on both rule difficulty and the salience of the cue for that rule. When the rule was relatively easy to use and its cue was salient, the participants adopted a rule-based strategy. In contrast, when the cue was not salient, even though the rule should have been equally easy to use, most participants eventually reported using a memory-based strategy. An ANOVA restricted to the rule and memory strategies provided statistical support for these observations. The only significant effect involving condition was the three-way interaction of condition  $\times$  block  $\times$  strategy [ $F(58,957) = 2.11$ ,  $MS_e = 0.142$ ,  $p < .001$ ].

The proportion of correct responses was much lower (near chance, .425) for the guessing strategy than for the rule (.882) and memory (.872) strategies [ $F(2,66) = 117.32$ ,  $MS_e = 0.031$ ,  $p < .001$ ]. Correct RTs were shorter when the participants reported using the rule strategy than when they reported using the memory strategy for both the first-letter artificial task (rule, 2.997 sec; memory, 3.132 sec) and the natural task (rule, 2.904 sec; memory, 3.108 sec), but they showed the opposite pattern for the last-letter artificial task (rule, 3.232 sec; memory, 3.174 sec). Again, RTs were clearly related to the dominant strategy in each task at the end of practice. The participants were faster on the dominant strategy than on the less preferred strategy. An ANOVA using blocks as a random effect provided statistical support for the observed patterns in terms of a significant two-way interaction of condition  $\times$  strategy [ $F(2,33) = 3.29$ ,  $MS_e = 0.034$ ,  $p < .05$ ], as well as a significant main effect of condition [ $F(2,33) = 7.90$ ,  $MS_e = 0.031$ ,  $p < .01$ ].

**Responses to novel stimuli.** For the first-letter artificial and natural conditions, the participants were most

likely to report using a rule (first-letter artificial = .504, natural = .496) followed by guessing (first-letter artificial = .388, natural = .342) on the novel trials. In the last-letter artificial condition, however, the participants reported guessing (.475) more than rule use (.371). Memory and other responses were relatively infrequent in all conditions. In an analysis of the proportion of times the guess, memory, and rule strategies were used as a function of condition for both training and novel items, there was only one statistically significant effect: the interaction of stimulus type and strategy [ $F(2,66) = 42.87$ ,  $MS_e = 0.037$ ,  $p < .001$ ]. As shown in Table 3, the participants guessed much more frequently on novel stimuli than on training stimuli. In contrast, the participants reported a memory strategy much more frequently with training stimuli than with novel stimuli, as would be expected, because each of the novel stimuli was shown only once, such that memory for training instances should have been ineffective as a strategy.

## Discussion

In Experiment 2, we used a first-letter artificial task that was equivalent to the natural task in the sense that the participants were required in both cases to classify stimuli on the basis of the vowel/consonant distinction of the first letter of a string. In both the natural and first-letter artificial conditions, performance at the end of the training session was largely rule based, as opposed to memory based. Because performance was virtually the same in the natural and first-letter artificial conditions, support was provided for the argument that the participants focused on the initial letter in both tasks. The rule difficulty explanation of the differences observed earlier between the natural task and



**Table 3**  
**Proportion of Use of the Guess, Rule, and Memory Strategies on Training and Novel Items in Experiment 2**

Stimulus Type	Guess	Rule	Memory
Training	.14	.40	.44
Novel	.40	.46	.12

the alphabetic versions of the artificial task is consistent with these results.

The results identify cue salience as well as rule difficulty as determiners of performance. The rule was the same in all three conditions of Experiment 2, but, in the last-letter artificial case, the cue letter was at the end of the word, rather than at the beginning, as it was in the natural and first-letter artificial conditions. Because scanning a linear array requires an increment of time (Lassaline & Logan, 1993), RTs were slower in the last-letter artificial condition, relative to those of the other two conditions. The decrease in cue salience not only caused responding to slow down overall in the last-letter artificial task relative to the other tasks, but it also caused a change in the preferred strategy. The rule strategy was preferred in the first-letter artificial and natural tasks, but the memory strategy was preferred in the last-letter artificial task, despite the fact that the one-letter rule was the same in all three cases. Thus, it is not only difficulty of the rule that affects strategy choice and strategy shift affordance but also the salience of the relevant cue to that rule.

## GENERAL DISCUSSION

Touron and Hertzog (2004) suggested that a useful way to address the issue of strategy shifts in cognitive skill acquisition is to manipulate the processing demands of a task so as to vary the cost/benefit ratio of feasible strategies. Indeed, from results obtained in a simulated air traffic control task, Schunn and Reder (2001) have argued that participants can, with practice, adjust their strategy to task demands, thus maximizing performance benefits. One manipulation that should affect processing demands in a classification task is rule difficulty. In Experiment 1 of the present study, as in earlier experiments (Bourne et al., 1999), we have shown that the strategy that participants eventually came to prefer, after extended practice, differed for binary classification tasks governed by a simple natural rule and a more difficult arbitrary alphabetic rule. When the rule governing category responses is relatively difficult (artificial hard), participants shift from an initial rule-based strategy to a memory-based strategy. When the rule is simple, however, providing a basis for rapid correct responding (natural task), participants do not shift strategies but stick with the initially adopted rule-based strategy. Likewise, making the artificial alphabetic rule easier should reduce the affordance of a shift from rule to memory, which it did in Experiment 1. The possibility that the participants were given insufficient exposure to task stimulus items in previous experiments (see Klapp et al., 1991) is ruled out by the fact that a second training

session in Experiment 1 produced performance functions in all three conditions essentially identical to those in the first session.

But the natural and artificial tasks used in Experiment 1 differed in a number of ways other than the defining rule. Most importantly, the stimuli in the natural task were familiar nouns and adjectives, whereas those in the artificial tasks were arbitrary letter strings. Furthermore, the rule for classifying stimuli in the natural task was based on the first letter—vowel or consonant—of the stimulus word, whereas classification in the artificial tasks required consideration of all letters in a stimulus. Experiment 2 showed that performance in an artificial task that allowed classification responses to be based on the first letter only was essentially the same in all respects, including strategy choices, as performance in the natural task. Thus, when classification can be based on a simple first-letter rule, use of that rule persists through extensive repeated exposure to task stimuli, and there is no evidence of a shift to memory-based performance. These results are contrary to the skill acquisition models of Logan (1988) and Rickard (1997) that expect, with sufficient practice and exposure to task stimuli, an eventual transition to memory-based performance under any conditions.

Touron and Hertzog (2004) interpret the rule-to-memory strategy shift in terms of the shift affordance of a given task. Experimental manipulations that increase the relative benefit of a memory strategy over a rule strategy are said to increase the task's affordance of a strategy shift. More difficult rules for classification create a greater affordance of a rule-to-memory shift. But, like Logan (1988) and Rickard (1997), Touron and Hertzog consider shifts in only one direction, rule to memory. The idea is that, if a task affords the use of a memory strategy to drive fluent processing and performance, skill acquisition is governed by a shift from a slow-acting rule strategy to a fast-acting memory strategy. No shift occurs if the rule-based process is sufficiently fast to preclude that shift. Our results suggest, furthermore, that strategy shifts can occur in either direction. When we introduced a 1-week retention interval between sessions, even the participants who acquired and came to prefer a memory-based strategy in Session 1 shifted to a rule-based strategy after the retention interval in Session 2. Apparently, the opportunity for forgetting is another manipulation that affects strategy use—in this case, affording a shift from memory-based performance to rule-based performance. In our view, either strategy (rule or memory) might have the better cost/benefit ratio, depending on task conditions and task demands and thus allowing for the possibility of a shift in either direction. It is important to note that, if this is the case, then the implication is that both strategies have costs associated with them. Thus, a memory-based strategy, at least in some circumstances, is not cost free or fully automatic, as the Logan model assumes.

Rule difficulty affects the cost/benefit ratio of a shift from rule-based to memory-based performance. Interpolation of a retention interval has the opposite effect, increasing the probability of a shift from a memory to a rule

strategy. Cue salience—that is, the location of the critical letter in a letter string stimulus—is still another variable that influences the cost/benefit ratio of a strategy shift. Assuming the typical left-to-right reading procedure, placing the critical letter in the last, as opposed to the first, position in a stimulus string requires the performer to scan the stimulus, adding a cost to the use of a rule-based classification strategy. Furthermore, costs are added only to the rule-based, not the memory-based, strategy. Thus, the cue salience is one more task manipulation increasing the affordance of a shift from rule to memory.

The method used to identify the participants' strategies in these experiments was based on self-report strategy probes, given immediately after each classification response. This report procedure was patterned after methods used by Tulving (1985) and others (e.g., Conway, Gardiner, Perfect, Anderson, & Cohen, 1997) in memory tasks to study the distinction between remembering (which corresponds to the memory strategy here) versus knowing (which is analogous to the rule strategy in the present task). This probe procedure has also been used with great success in studies of mental arithmetic (Delaney, Reder, Staszewski, & Ritter, 1998) and alphabet arithmetic (Rickard, 1994, 2004), as well as binary classification (Bourne et al., 1999). However, there appear to be some possible difficulties with the interpretation of results obtained by this procedure. For example, rather than reflecting true strategy transitions, these reports might merely be correlated with performance changes over practice. Alternatively, transitions might reflect participants' attempts to comply with perceived task demands, as performance improves. Finally, it is possible that probes cause participants to adjust their speed-accuracy criterion or induce strategy shifts when they might not otherwise occur—an effect called *reactivity* (Rickard, 2004).

There have been a few attempts to directly assess the validity of the probe procedure, and the outcome of these attempts has generally been supportive. Green, Cerella, and Hoyer (2000) collected strategy probes from only half of their participants and found no difference between groups, suggesting the absence of reactivity attributable to the probes. Rickard (2004) showed that probed participants shifted slightly, but significantly, earlier in the trial sequence than did nonprobed participants, suggesting some reactivity. He demonstrated, however, by item analysis that, in his alphabet arithmetic task where the strategy transition is always in the direction of algorithm to memory retrieval, reports of the algorithm strategy were exclusively observed in the slow cluster of items and memory reports were exclusively observed in the faster cluster.

Our results, using a different task, are entirely consistent with Rickard's (2004) outcome. However, we are able to offer additional evidence in support of the validity of this procedure on the basis of performance on novel items that were presented at regular intervals during training. If the report procedure is valid, then the most frequent strategy on novel items should be guess (because these items have never been seen before) or rule (if the participant has dis-

covered the basis of the binary classification). The memory strategy should be reported for novel items only when they are confusable with training items (Allen & Brooks, 1991; Johansen & Palmeri, 2002). This outcome is, in fact, exactly what we observed. Overall, these results suggest that the strategy probe procedure has acceptable validity, although there is still a possibility of some reactivity.

In addition to its simplicity and directness, the self-report procedure has one other possible advantage over techniques based on an examination of critical transfer items. As discussed by Blair and Homa (2003) and as we found in our examination of novel items, the incidence of rule use might be high on transfer trials even when a memory-based strategy is dominant on training items. Therefore, it is possible that any measure of strategy use based only on generalization to transfer items may give a distorted picture of the relative dominance of competing strategies.

In conclusion, it should be noted that the distinction between rule-based and memory-based (exemplar) performance is not limited to simple binary classification tasks and associated models. In fact, this distinction has been discussed in such diverse contexts as alphabet arithmetic (e.g., Logan, 1988; Rickard, 1997), memory (e.g., Gardiner, Gawlik, & Richardson-Klavehn, 1994), categorization (e.g., Johansen & Palmeri, 2002), and artificial language learning (e.g., Reber, Allen, & Regan, 1985). Our findings suggest that participants in any of these tasks are likely to base their responses on either rules or memory at different points of practice, depending on specific task affordances. Affordances, in turn, depend on such variables as rule difficulty and cue salience. It is a challenge to future theoretical developments in these diverse domains to accommodate the notion of task affordances that lead to particular strategies and shifts between strategies.

## REFERENCES

- ALLEN, S. W., & BROOKS, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General*, **120**, 3-19.
- BLAIR, M., & HOMA, D. (2003). As easy to memorize as they are to classify: The 5-4 categories and the category advantage. *Memory & Cognition*, **31**, 1293-1301.
- BOURNE, L. E., JR., HEALY, A. F., PARKER, J. T., & RICKARD, T. C. (1999). The strategic basis of performance in binary classification tasks: Strategy choices and strategy transitions. *Journal of Memory & Language*, **41**, 223-252.
- CONWAY, M. A., GARDINER, J. M., PERFECT, T. J., ANDERSON, S. J., & COHEN, G. M. (1997). Changes in memory awareness during learning: The acquisition of knowledge by psychology undergraduates. *Journal of Experimental Psychology: General*, **126**, 393-413.
- DELANEY, P. F., REDER, L. M., STASZEWSKI, J. J., & RITTER, F. E. (1998). The strategy-specific nature of improvement: The power law applies by strategy within task. *Psychological Science*, **9**, 1-7.
- GARDINER, J. M., GAWLIK, B., & RICHARDSON-KLAVEHN, A. (1994). Maintenance rehearsal affects knowing, not remembering; elaborative rehearsal affects remembering, not knowing. *Psychonomic Bulletin & Review*, **1**, 107-110.
- GREEN, D. R., CERELLA, J., & HOYER, W. J. (2000, November). *Do strategy probes affect the probability of item retrieval?* Poster presented at the 41st Annual Meeting of the Psychonomic Society, New Orleans.
- HEALY, A. F., KING, C. L., CLAWSON, D. M., SINCLAIR, G. P., RICKARD, T. C., CRUTCHER, R. J., ET AL. (1995). Optimizing the long-term retention of skills. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Learning*

- and memory of knowledge and skills: *Durability and specificity* (pp. 1-29). Thousand Oaks, CA: Sage.
- JOHANSEN, M. K., & PALMERI, T. J. (2002). Are there representational shifts during category learning? *Cognitive Psychology*, **45**, 482-553.
- KLAPP, S. T., BOCHES, C. A., TRABERT, M. L., & LOGAN, G. D. (1991). Automatizing alphabet arithmetic: II. Are there practice effects after automaticity is achieved? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 196-209.
- LASSALINE, M. E., & LOGAN, G. D. (1993). Memory-based automaticity in the discrimination of visual numerosity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 561-581.
- LOGAN, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492-527.
- LOGAN, G. D., & KLAPP, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 179-195.
- RAYMOND, W. D., FISHER, J. A., & HEALY, A. F. (2002). Linguistic knowledge and language performance in English article variant preference. *Language & Cognitive Processes*, **17**, 613-662.
- REBER, A. R., ALLEN, R., & REGAN, S. (1985). Syntactic learning and judgment, still unconscious and still abstract: Comment on Dulany, Carlson, and Dewey. *Journal of Experimental Psychology: General*, **114**, 17-24.
- RICKARD, T. C. (1994). *Bending the power law: The transition from algorithm-based to memory-based performance*. Unpublished doctoral dissertation, University of Colorado, Boulder.
- RICKARD, T. C. (1997). Bending the power law: A CMPL model of strategy shifts and the automatization of cognitive skills. *Journal of Experimental Psychology: General*, **126**, 288-311.
- RICKARD, T. C. (2004). Strategy execution in cognitive skill learning: An item-level test of candidate models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 65-82.
- RICKARD, T. C., & BOURNE, L. E., JR. (1995). An identical-elements model of basic arithmetic skills. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Learning and memory of knowledge and skills: Durability and specificity* (pp. 255-281). Thousand Oaks, CA: Sage.
- SCHUNN, C. D., & REDER, L. M. (2001). Another source of individual differences: Strategy adaptivity to changing rates of success. *Journal of Experimental Psychology: General*, **130**, 59-76.
- TOURON, D. R., & HERTZOG, C. (2004). Strategy shift affordance and strategy choice in young and old adults. *Memory & Cognition*, **32**, 298-310.
- TULVING, E. (1985). Memory and consciousness. *Canadian Psychologist*, **26**, 1-12.

(Manuscript received June 7, 2004;  
revision accepted for publication May 16, 2005.)