Memory & Cognition
2001, 29 (2), 254-266

Awareness and working memory in strategy adaptivity

CHRISTIAN D. SCHUNN
George Mason University, Fairfax, Virginia

and

MARSHA C. LOVETT and LYNNE M. REDER
Carnegie Mellon University, Pittsburgh, Pennsylvania

To further the understanding of the mechanisms of strategy choice, in three experiments, we investigate the role of explicit awareness and working memory in strategy adaptivity. Experiment 1 provided correlational evidence that individual differences in strategy adaptivity to changing base rates are related to individual differences in awareness of those changes but appear not to be related to individual differences in working memory capacity. Experiment 2 replicated the role of awareness, and the results suggest that awareness at the time of the base-rate change, rather than afterwards, is related to increased strategy adaptivity. Experiment 3 measured working memory capacity using a different procedure and manipulated working memory load with a dual-task procedure; again, no apparent role of working memory capacity in strategy adaptivity was found. This juxtaposition of findings presents a challenge for existing models of strategy choice.

The world is an ever-changing place, and tasks that we perform repeatedly frequently change in their characteristics. How do we adapt to such a changing world? One method is by changing our strategy for performing the task at hand. As aspects of a task change, the optimal strategy often changes, so adapting one's strategies is helpful. For example, using the World-Wide Web to answer a question is becoming an increasingly successful strategy in many domains.

There is an important twist in this characterization of how people adapt through changing strategies: Any given person almost never switches from using only one single strategy A to using only a different single strategy B. Instead, examinations of performance in a wide range of domains have shown that almost everyone uses many different strategies for a given task at a given point in time (Reder, 1982; Siegler, 1996). Thus, when people adapt to changing task characteristics, there is change in the *distribution* of how they use their strategies. For example, several years ago, a person searching for a new phone number might have used the strategy of calling directory assistance 50% of the time and using a phone book 50% of the time, whereas today that same person might call directory assistance 20% of the time, use a phone book 30% of the time, and use http://bigfoot.com 50% of the time. The important consequence of multiple-strategy use is that a model of changing strategy selections must be imbedded within a model of distributional strategy selection.

Researchers have proposed several different models of strategy choice and strategy change (Anderson & Lebiere, 1998; Lovett & Anderson, 1996; Reder, 1982, 1987, 1988; Reder & Schunn, 1996; Siegler & Shipley, 1995; Siegler & Shrager, 1984). Although there are significant differences among these models, there is an important commonality: Each model assumes that people keep track of the base rates of success of the different strategies and prefer the strategies that have higher success base rates (similar to Thorndike's, 1913, law of effect). This account of processing implies that, as the task characteristics change, the individual will experience different success base rates for each strategy and so learn to prefer different strategies.

Despite the existence of several models of strategy choice, many basic questions remain about the mechanism by which people adapt their strategy use in response to changing success base rates. This paper attempts to address two of these basic questions. One question concerns where the information about strategy success rates is stored: explicitly in working memory or in some more implicit long-term storage? Most models of strategy choice do not describe where the information is stored. However, since many of these models posit no decay or interference to the base-rate information (Lovett & Anderson, 1996; Siegler & Shipley, 1995; Siegler & Shrager, 1984), one could reasonably conclude that the information was assumed to be in long-term storage rather than in

This work was supported by Grant N00014-95-1-0223 to L.M.R. from the Office of Naval Research through Helen Gigley. A preliminary form of Experiment 1 is reported in Schunn and Reder (1998). We thank Mark Johnson, Rebecca Stotzer, Jay Stearns, Catherine Duda, and Scott Filippino for help in collecting the data, and we thank Pat Kyllonen for making the CAM4 available to us. Correspondence should be addressed either to C. D. Shunn, Department of Psychology, MSN 3F5, George Mason University, Fairfax, VA 22030-4444 (e-mail: schunn@gmu.edu), or to M. C. Lovett and L. M. Reder, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213 (e-mail: lovett@cmu or reder@cmu.edu).

working memory. In the models of choice that do include decay (Anderson & Lebiere, 1998), the decay is quite slow and the base-rate information (that directly influences strategy choices) is not stored in working memory; rather, it is assumed that the information is stored directly in with the long-term representation of the strategies (i.e., the production rules). The second question concerns whether adapting to changing base rates requires explicit awareness: Does it involve modifying implicit parameters in a gradual implicit learning process, or does it involve conscious effort triggered by noticing a change? Again, most models of strategy choice do not describe the accessibility of the base-rate information. One exception is the Lovett and Anderson (1996) account, which places the success base-rate information as parameters associated with production rules. According to this model, there is no mechanism by which the system can directly access or reason about these parameter settings. Thus, they are not accessible to introspection and so must be considered implicit. Consistent with the Lovett account, Reder and Schunn (1996) argued that metacognitive control (i.e., strategy choice) was driven entirely by implicit memory and implicit learning (see also Lemaire & Reder, 1999; Spehn & Reder, 2000).

This implicit control view is directly opposed by a number of accounts of metacognition (Davidson, Deuser, & Sternberg, 1994; Metcalfe, 1994; Nelson & Narens, 1994). These accounts assume that we have metacognitive awareness so that we can control and change our behavior. For example, Nelson and Narens (1994) describe people as "systems containing self-reflective mechanisms for evaluating (and re-evaluating) their progress and for changing their on-going processing" (p. 7).

The two questions are potentially related. One simple and common view is that if adaptivity to changing base rates requires explicit awareness, then it is likely that the base-rate information is stored in working memory, at least while it is being processed. Similarly, if the information is stored in working memory, then it is more likely to be available to explicit processing. Conversely, if the information is not in working memory, then it is likely to be processed implicitly, and if it is processed implicitly, then it not likely to rely on working memory. That is, a common assumption is that conscious access requires working memory and that the contents of working memory are freely available to consciousness (see Ericsson & Simon, 1993).

The goal of this paper is to examine these two questions. More specifically, we will examine the impact of working memory capacity and explicit awareness of base-rate shifts on strategy adaptivity to base-rate shifts. Previous work by Schunn and Reder (1998, in press) bears directly on these questions. They examined strategy adaptivity in the context of a complex simulated Air Traffic Control task. They found that individual differences in strategy adaptivity to base-rate shifts were most strongly related to individual differences in inductive reasoning ability. They argued that inductive reasoning may have been required for noticing the pattern of the base-rate change. This account seems to invoke a role for explicit awareness of base-rate changes in adaptivity to those changes. However, explicit awareness was never directly examined in those studies. In the present experiments, we examined explicit awareness directly.

Schunn and Reder (1998, in press) also found that individual differences in working memory capacity were positively (but not as strongly) related to adaptivity. However, since working memory capacity is quite strongly correlated with inductive reasoning ability (Carpenter, Just, & Shell, 1990), it is unclear whether the relationship between adaptivity and working memory capacity was mediated through inductive reasoning ability. In the present experiments, we examined the role of working memory capacity in strategy adaptivity in further depth.

THE BUILDING STICKS TASK

Before turning to the experiments, we will introduce the problem-solving task that was used in all of the experiments. The building sticks task (BST; Lovett & Anderson, 1996) is a problem-solving task similar to the classic water jars task (Luchins & Luchins, 1950). For a given BST problem, participants must add and subtract an unlimited supply of three different-sized building sticks to create a stick of the desired length (see Figure 1, *initial state*). BST problems can be solved by one of two strategies. The *undershoot* strategy involves starting with a building stick that is shorter than the desired stick and then lengthening that stick by additional building stick lengths until the desired stick's length is reached. In contrast, the *overshoot* strategy involves starting with the building stick that is longer than the desired stick and then shortening that stick by the other building stick lengths. As Figure 1 shows, participants choose between these two strategies in their first step.

For example, suppose the desired stick was of length 14 units, and the three sticks, a, b, and c, were of lengths 2, 17, and 10, respectively (note that participants were never given the exact numerical lengths of the sticks and had to visually estimate the length of each stick). To obtain the desired stick length of 14 units, the participants might start with stick b of 17 units and remove segments (the overshoot strategy), or the participants might start with stick c of 10 units and add more segments (the undershoot strategy). In this example, a solution can be obtained only by the undershoot strategy ($c + a + a = 10 + 2 + 2 = 14$). The overshoot strategy will not work because subtracting lengths a and c from b will never lead exactly to a stick of length 14 units (because, in this case, b is odd, and a, c, and the goal stick are all even). Of course, in other problems, the overshoot strategy may be the correct one to use.

Within this task, it is easy to manipulate the base rates of success of the undershoot and overshoot strategies. Each problem can be designed to be solvable by either undershoot or overshoot (but not both), and then the propor-

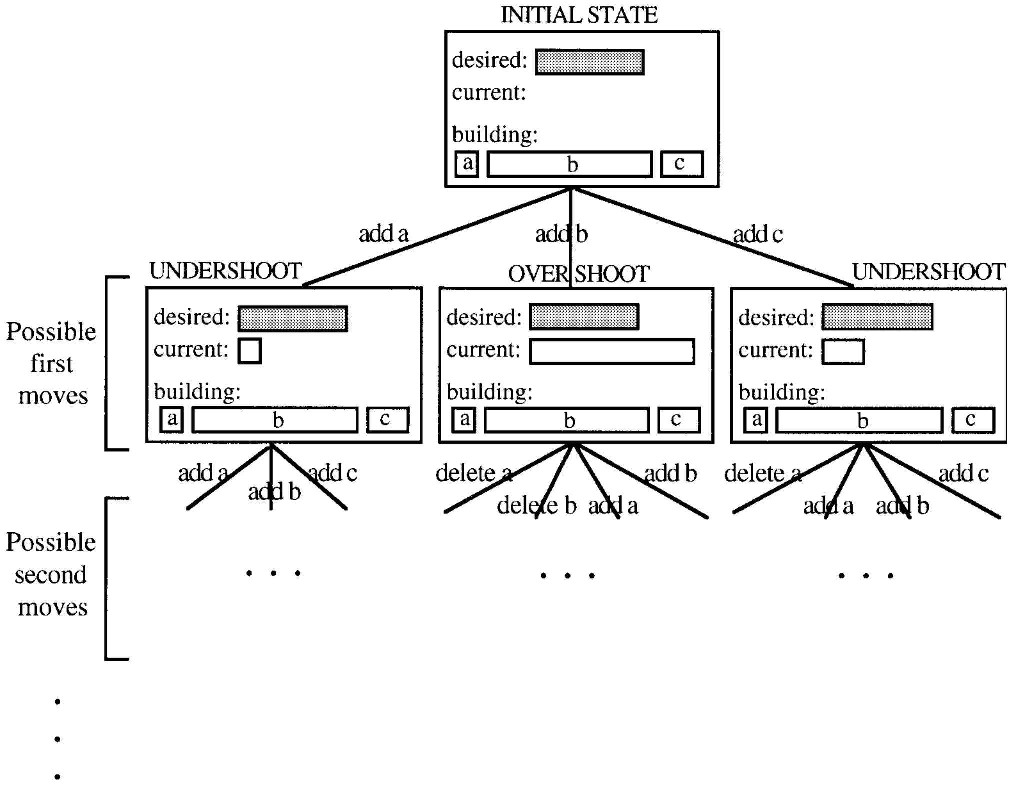

Figure 1. Initial and subsequent states in a BST problem. The participant's task is to build a current stick (initially length 0) that matches the desired stick in length by adding and subtracting various combinations of the building stick lengths. From "History of Success and Current Context in Problem Solving: Combined Influences on Operator Selection," by M. C. Lovett and J. R. Anderson, 1996, *Cognitive Psychology*, *31*, p. 175. Copyright 1996 by Academic Press. Reprinted with permission.

tion of problems with each solution type can be varied across blocks of time. In this way, it is possible to directly control the success base rates of each strategy and thereby clearly measure individual differences in adaptivity to change success rates.

The basic template of the present experiments is as follows. Participants are given a large number (60–70) of BST problems to solve, during which time the base rate of success of the overshoot and undershoot problems is manipulated. Strategy adaptivity is defined as the amount of change in overshoot strategy use in response to those changing success base rates. Experiment 1 examined the roles of both explicit awareness and individual differences in working memory capacity and inductive reasoning ability in strategy adaptivity. Experiment 2 followed up on the findings on the role of explicit awareness. Experiment 3 followed up on the role of working memory capacity.

EXPERIMENT 1

Method

Participants. Fifty-six Carnegie Mellon University (CMU) undergraduates participated for course credit and were randomly assigned to the two conditions (50–80–20 or 50–20–80, to be described below). Because of a software bug, the participants were three times as likely to be assigned to the 50–80–20 condition, producing 42 participants in the 50–80–20 condition and 14 in the 50–20–80 condition.

Procedure. The participants were first given several ability tests. The working memory test was the Synthesis Add Matrices test taken from the CAM4 (Kyllonen, 1993, 1994, 1995). This test was designed to be a measure of spatial working memory. There were also three short inductive spatial tests taken from the CAM4: Figure Sets, Figure Series, and Figure Matrices. The inductive reasoning tests were included to examine whether any correlations between adaptivity and working memory capacity are mediated through inductive reasoning ability. The mean score across all three tests was used as a measure of inductive reasoning ability. See Appendix A for a description of the working memory and inductive reasoning tests.

Following the ability tests, the participants were given the BST. Each participant was given 70 BST problems to solve. The problems, although differing in absolute lengths (i.e., there were no duplicates), were all "neutral" looking problems (i.e., undershoot and overshoot appeared to bring one equally close to the goal). Auditory feedback was used to signal to the participants whether they had solved a problem within the desired five moves.

To measure strategy adaptivity, the base rates of success of the undershoot and overshoot strategies were manipulated over time. For 10 trials, both strategies were equally successful (i.e., 5 solved by overshoot and 5 solved by undershoot). For the next 30 trials, one strategy was successful on 80% of trials. For the final 30 trials, the other strategy was successful on 80% of trials. Varied across two

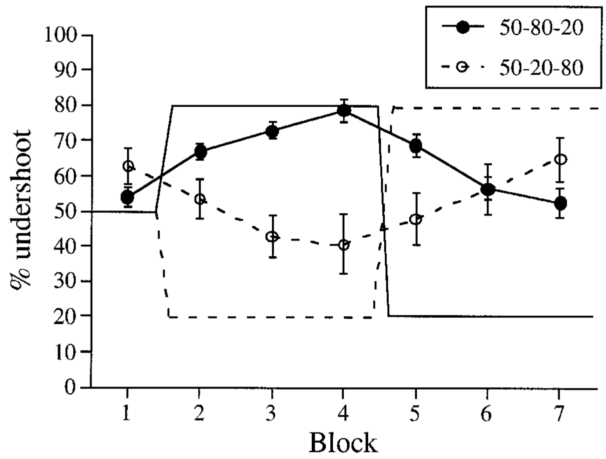

Figure 2. Mean (and *SE*) use of the undershoot strategy within each block in each condition of Experiment 1, along with the manipulated base rates.

conditions was which strategy was more successful first. That is, in one condition (50–80–20), the undershoot strategy was successful on 50%, then 80%, then 20% of trials; in the other condition (50–20–80), undershoot was successful on 50%, then 20%, then 80% of trials. Strategy adaptivity was defined as the change in mean strategy use from Block 4 (the end of the second phase) to Block 7 (the end of the third phase).

To measure explicit awareness of base-rate change, the participants were asked a series of questions at the end of the experiment. Specifically, the overshoot and undershoot strategies were described to the participants, and the participants were asked, "Did there seem to be any pattern to which strategy worked?" and "Did you notice any changes in the effectiveness of one strategy or the other as the experiment progressed?" The participants' responses were coded for awareness of the base-rate manipulation: 1 was given if the participant reported awareness of a change and correctly described the direction of the change, and 0 was given otherwise. Only a few participants had intermediate awareness: apparent awareness of change, but inaccuracy in the reported direction. These participants behaved like the unaware participants and thus were pooled into that group. Sixty percent of the data was recoded by a second coder, and the interrater reliability was 97%.

Results

Overall adaptivity. A participant's first choice on each trial was used to categorize that trial's strategy use as overshoot or undershoot. The 70 trials were divided into blocks of 10 trials. Figure 2 illustrates the mean undershoot use within each block within each condition, as well as the manipulated success rates of undershoot within each block for each condition. The 50–80–20 condition showed the expected increase in use of the undershoot strategies in Blocks 2–4 and then the expected decrease in use of the undershoot strategies in Blocks 5–7. The 50–20–80 condition showed the reverse pattern. Thus, on average, the participants adapted to the base-rate changes. Because the two conditions behaved so similarly, the two conditions were pooled (to gain greater power for the individual difference analyses) by reversing the values for the 50–20–80 condition (i.e., subtracting the proportions from 1).

Individual differences. Are there large individual differences in strategy adaptivity? Using a Monte Carlo simulation (N = 1,000), we can establish the expected variation among participants assuming they all had true probabilities of selecting undershoot in a block that corresponded to the mean undershoot use across participants in each block. Because there were only 10 trials per block and each trial produced a binary outcome (undershoot or overshoot), one would expect a certain amount of variability due simply to sampling noise. Figure 3 plots the observed frequency histogram of the participants' adaptivity (as measured by the difference between an individual's mean strategy use in Blocks 4 and 7) and the expected distribution from the Monte Carlo simulation.

The observed distribution is flatter than the expected distribution, indicating more variation among participants than could be attributed to sampling noise. For example, we would expect 9% of the participants to have zero or less adaptivity by chance; 20% of the participants actually fell into this group. At the other extreme, we would expect only 8% of the participants to have greater than .5 adaptivity by chance; in fact, 20% of the participants displayed this higher level of adaptivity. The observed distribution of adaptivity differed statistically from the expected distribution [$\chi^2(df = 3; N = 56) = 13.0, p < .01$].[1] These comparisons establish that there are individual differences in strategy adaptivity that are not attributable to chance variation. In subsequent analyses, the participants whose adaptivity (as defined above) was at or below zero will be called the *nonadaptive* participants.

Is it possible that these individual differences in adaptivity are due to floor or ceiling effects in strategy use (i.e., always using undershoot or never using undershoot)? Examinations of the distribution of mean percent undershoot across the 70 trials revealed that none of the 11 nonadaptive participants was outside of the 20%–80% range, and only 2 were outside of the 25%–75% range. Therefore, most of the nonadaptive participants did use both overshoot and undershoot strategies regularly, just not adaptively over time.

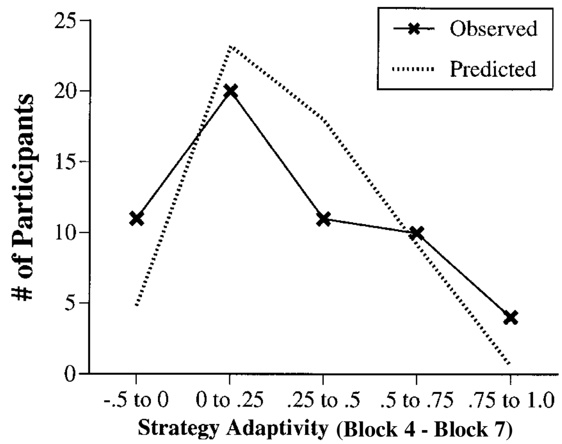

Figure 3. Observed and predicted frequency histogram of strategy adaptivity levels in Experiment 1.

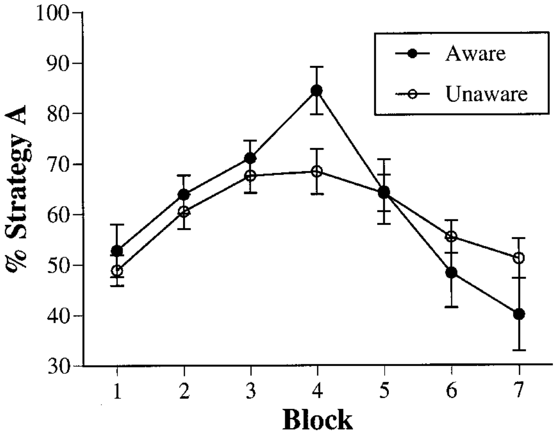

Figure 4. Mean (and *SE*) strategy use for the aware and unaware participants within each block of Experiment 1.

Another possible explanation for why some participants may not have adapted is that they were simply doing the task too quickly to notice the change in base rates. To investigate this possibility, a block $\times$ adaptivity (adaptive/nonadaptive) analysis of variance (ANOVA) was conducted on the mean time to make the first move (i.e., click on the desired first stick and click on its target location). There was a speed-up over blocks [$F(6,324) = 7.1$, $MS_e = 1.5$, $p < .0001$], as the mean time changed from 5.7 sec in Block 1 to 4.5 sec in Block 7. However, the effect of adaptivity was only marginally significant [$F(1,54) = 2.1$, $MS_e = 36.7$, $p < .15$]. If anything, the adaptive participants were slightly faster (4.5 sec vs. 5.6 sec; and they maintained this speed advantage in all 7 blocks, although the difference decreased slightly in size over blocks) [$F(6,324) = 1.9$, $MS_e = 1.5$, $p < .1$]. Therefore, the nonadaptive participants were not simply choosing too quickly: 5.6 sec is a long time to make two mouse clicks.

Awareness and adaptivity. How did explicit awareness of the base-rate changes (as measured by the debriefing questions) relate to the strategy adaptivity? As described earlier, the participants were divided into those who explicitly noticed the direction of the shift (aware, $N = 18$) and those who did not (unaware, $N = 37$). An awareness $\times$ block ANOVA conducted on the mean strategy use in each block revealed a significant interaction [$F(6,318) = 3.6$, $MS_e = 0.02$, $p < .005$]. The unaware participants did adapt, but the aware participants made the transition more precipitously (see Figure 4). Focusing on the Block 4 to Block 7 transition, the aware participants showed a greater transition than the unaware participants (means of .44 and .17, respectively) [$F(1,53) = 12.7$, $MS_e = 0.07$, $p < .001$].

Working memory, inductive reasoning, awareness, and adaptivity. Were the psychometric tests predictive of adaptivity or explicit awareness? Table 1 presents the correlation matrix of working memory ability, inductive reasoning, awareness, and strategy adaptivity. Working memory and inductive reasoning appeared to be unrelated to amount of adaptivity and negatively related to explicit awareness of the shift. The relationship between working memory and awareness appeared to be mediated through inductive reasoning: (1) working memory and inductive reasoning correlate quite well with one another; (2) inductive reasoning contributes significant additional variance when working memory is entered first into a regression on awareness; and (3) working memory does not contribute significant additional variance when inductive reasoning is entered first. Thus, the overall model appears to be that working memory is positively related to inductive reasoning, inductive reasoning is negatively predictive of awareness, and awareness is positively predictive of adaptivity; there is no direct role of working memory in awareness or adaptivity.

Discussion

As found in other domains (Reder & Schunn, 1999; Schunn & Reder, 1998, in press), there are meaningful individual differences in adaptivity in the BST domain that appeared not to be attributable to various artifacts. Thus, even when strategy success rates are carefully controlled across participants, participants appear to differ significantly in their abilities to change strategy use in response to shifting base rates of success.

This experiment provided a suggestion that the individual differences may be related to explicit awareness of base-rate change: Awareness is associated with faster or larger changes. Of course, the present experiment did not establish the causality of awareness in adaptivity. Because the relationship found thus far is only correlational, all three logical interpretations remain as possibilities: (1) explicit awareness leads to greater shifts in strategy use; (2) greater shifts in strategy use are more likely to lead to explicit awareness after the fact; and (3) some third factor (e.g., working memory capacity or inductive reasoning skill) leads to both greater shifts in strategy use and a greater likelihood of noticing the base-rate change. The relationship between inductive reasoning and awareness and the lack of a direct relationship between adaptivity and either working memory capacity or inductive reasoning ability provide some evidence against the third explanation. In Experiment 2, we attempted to provide more information about the timing of the relationship be-

Table 1
Correlations Between Working Memory Spatial, Inductive Reasoning Spatial (Mean Across Three Tests), Explicit Awareness of the Base-Rate Shift (Yes = 1, No = 0), and Amount of Strategy Adaptivity (Block 4 Minus Block 7) in Experiment 1

	Working Memory	Inductive Reasoning	Aware of Shift?
Inductive reasoning	.45		
Aware of shift?	−.27	−.33	
Adapt amount	−.06	−.08	.44

Note—Correlations with $|r| > .26$ and $|r| > .35$ are significant at $p < .05$ and $p < .01$, respectively.

tween adaptivity and awareness, thus providing relevant information regarding the causality of the relationship.

A second contribution of Experiment 1 was that it provided some evidence against a direct relationship between working memory and strategy adaptivity—in other words, strategy selection appears to not require storing detailed information about strategy successes and failures in working memory. Of course, one must be careful about conclusions drawn from null results, especially in a correlational study with $N = 56$. The 95% confidence interval for the correlation between working memory and adaptivity is $(-.32, .21)$. Thus, it is possible that there is actually a small positive correlation. In Experiment 3, we reexamined this relationship by using a different estimator of working memory capacity and by manipulating working memory load with a secondary task.

The negative correlation between awareness and inductive reasoning deserves further comment. Why would apparently more intelligent people seem to have performed worse in the task? In resolving this puzzle, there are several additional pieces needed. First, this task was quite difficult overall, and there is no trick that will guarantee very good performance. Overall, the participants had solved the problems in fewer than five moves for only 51% of the problems. In this kind of situation, the more intelligent participants may have been busy trying to figure out a good strategy in a situation in which there was no good strategy and thus may have missed the relatively simple strategy of just following the base rates. Alternatively, the more intelligent participants might have become bored by this task in which all the problems essentially look alike and there is no clear basis for picking one strategy over another. For this reason, in Experiments 2 and 3, we used a set of problems that varied in terms of which strategy looked better. Finally, although adaptivity was positively related to overall task performance ($r = .47$, $p < .001$, with mean success rate across all blocks), inductive reasoning was weakly positively related to overall task performance ($r = .2$, $p < .15$). Thus, more intelligent people did not perform worse overall.

EXPERIMENT 2

The goal of Experiment 2 was to provide more information about the role that awareness plays in strategy adaptivity. In the posttask debriefing, participants were again asked whether they were aware of the base-rate shift, but this time they were also asked at what point they came to this awareness: in the middle of the task, toward the end of the task, or only by thinking back on their experiences in response to the question. If participants primarily became aware of the change toward the end of the experiment or in the debriefing, or if those participants who had a very late awareness were just as adaptive as those having an early awareness, then it is very unlikely that awareness actually produces the increases in adaptivity.

A second feature of Experiment 2 was that the problems varied more in their appearance than did those in Experiment 1 in a particular way: Rather than having all neutral problems, half the problems were biased toward overshoot and half the problems were biased toward undershoot. Previous research has shown that participants' BST strategy selections are very sensitive to the looks of the problem (Lovett & Anderson, 1996). It was possible to have the problem looks be predictive or unpredictive of which strategy is correct. Because both kinds of situations are interesting cases, this was manipulated across participants. For the participants in the predictive condition, the problems could generally, but not always, be solved using the strategy indicated by the problem looks. For the participants in the nonpredictive condition, the looks of the problem did not at all predict which strategy should be selected. Previous research has established that participants come to rely on the looks of the problem much less when they prove to be unpredictive (Lovett & Schunn, 1999). In such a situation, participants may rely more heavily on base-rate information and may be more aware of shifts in base rates.

Method

Participants. Forty-six CMU undergraduates participated for course credit and were randomly assigned in equal numbers to one of two conditions (predictive or nonpredictive, to be described below). Due to a computer problem, BST data were not collected from 1 participant in the predictive condition.

Procedure. The BST was the same as that in Experiment 1, with the following exceptions. Because the focus was on the transition between the strongly overshoot-biased block and the strongly undershoot-biased block, Experiment 2 used only 60 BST problems (i.e., deleted the first 10 problems at 50/50 success rates). Because all previous studies with the BST have found no difference between the ordering of undershoot first versus overshoot first, one fixed order was selected for all participants: Overshoot was 70% successful for the first 30 problems, and undershoot was 70% successful for the second 30 problems.

Within each block of 10 problems, half the problems were perceptually biased toward overshoot and half were biased toward undershoot. If the three sticks have lengths a, b, and c, and the goal stick has length g, this bias was operationalized as having $(b - g) - (g - c)$ in the range [45,55] pixels for undershoot biased problems and in the range $[-55, -45]$ pixels for overshoot biased problems. For example, a problem with building sticks of lengths 16, 256, and 60 and a goal length of 133 is biased toward undershoot, because the stick c brings one 50 pixels closer to the goal than does stick b. Again, problems did not repeat. In the predictive condition, the looks of the problem were 73% predictive in each block (i.e., problems could be solved by the strategy suggested by the perceptual cues 73% of the time). In the nonpredictive condition, the looks of the problem were 53% predictive in each block—it was impossible to achieve exactly 50% predictivity using 30 problems and success base rates of 30% or 70% overshoot.

Rather than assessing awareness of the base-rate changes with open-ended questions administered verbally, Experiment 2 used a written multiple-choice format precisely targeting awareness of base-rate changes in overshoot and undershoot proportions. Specifically, the participants were given a brief definition of the undershoot and overshoot strategies and were then asked whether the propor-

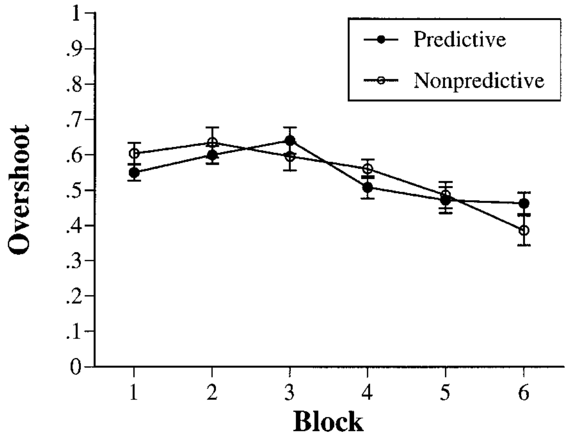

Figure 5. The mean proportion (and *SE*) of overshoot use within each block for each condition in Experiment 2.

tion of problems that could be solved by undershooting was higher, lower, or the same at the end of the experiment as at the beginning (three-alternative forced choice). Next, the participants were asked when they noticed this pattern: "just now by thinking back over my experience," "in the middle of the experiment," "towards the very end of the experiment," or "not applicable—there wasn't any pattern."

The participants were not given any psychometric tests in this experiment.

Results

Sensitivity to problem looks. In contrast to Experiment 1, all of the problems in Experiment 2 were biased such that either undershoot or overshoot looked like it would be the more successful strategy. The participants were quite sensitive to this aspect of the task. They usually selected overshoot when overshoot looked best (mean overshoot rate = .82) and usually selected undershoot when undershoot looked best (mean overshoot rate = .26). This bias held even when the looks were not predictive of which strategy would be correct (means of .83 and .25 in the predictive condition, and .81 and .26 in the nonpredictive condition).

Overall adaptivity. Figure 5 presents the mean proportion of overshoot strategy use within each block of 10 trials for each condition. The participants began the experiment with no bias toward either strategy and gradually developed a preference for the more successful strategy. As the figure demonstrates, there was a significant effect of block [$F(5,215) = 13.7$, $MS_e = 0.020$, $p < .0001$], no effect of condition (predictive/nonpredictive) [$F(1,43) < 1$], and a nonsignificant interaction [$F(5,215) = 1.7$, $MS_e = 0.02$, $p < .15$]. More specifically, there was no difference between the conditions in the size of the strategy shift between the third and sixth blocks (mean strategy shifts of .18 and .21 in the predictive and nonpredictive conditions, respectively) [$F(1,43) < 1$, $MS_e = 0.06$]. For the remaining analyses, we pooled across the predictive and nonpredictive conditions, since the behavior in the two conditions was quite similar and there were no significant interactions.

Awareness and adaptivity. On the basis of their responses to the awareness questionnaire, the participants were divided into four groups: those indicating the correct direction of change and reporting awareness of this pattern in the middle of the experiment (correct immediate, $n = 13$), those indicating the correct direction of change and reporting awareness at the end of the experiment or just when the question was asked (correct later, $n = 11$), those indicating the wrong direction of change (incorrect, $n = 10$), and those reporting that no change had occurred (unaware, $n = 11$).

Figure 6 presents the mean adaptivity (proportion overshoot use in Block 3 minus proportion overshoot use in Block 6) as a function of awareness. An ANOVA of awareness on adaptivity was only marginally significant [$F(3,41) = 2.1$, $MS_e = 0.052$, $p < .1$]. However, as Figure 6 reveals, there was considerable variability in the incorrect group, which likely reflected a mixture of participants who were actually aware of the correct change but misremembered or misresponded on the questionnaire and participants who were unaware of a change. Removing the incorrect participants, the overall ANOVA was statistically significant [$F(2,32) = 4.9$, $MS_e = 0.036$, $p < .02$]. Although all groups showed nonzero mean adaptivity, only the participants in the correct immediate group showed higher levels of adaptivity (Bonferroni–Dunn ps $< .05$, for both contrasts against correct later and unaware). The participants who became aware of the change only at the end of the experiment or when asked the awareness question showed no more adaptivity than those participants unaware of the change (Bonferroni–Dunn $p > .9$).

Discussion

In Experiment 2, awareness was again correlated with adaptivity, even when a more complex set of non-neutral problems was used. Thus, even when participants have relevant information to which they can attend, awareness of base-rate changes is an important correlate of adaptivity. More importantly, in Experiment 2, there was evidence that only immediate awareness was correlated with greater adaptivity. However, it is important to note that this evidence was only a correlation and did not establish a causal relationship between awareness and adaptivity.

EXPERIMENT 3

The goal of Experiment 3 was to further investigate the apparent lack of a relationship between working memory and adaptivity from Experiment 1. First, rather than simply examining the correlations between individual differences in capacity and individual differences in strategy adaptivity, we manipulated available working memory through a dual-task manipulation in Experiment 3. Second, we used a verbal working memory test in Experiment 3 (in contrast to the spatial working memory test of Experiment 1) and a different method for estimating individual differences in working memory capacity. Instead of administering a psychometric test and

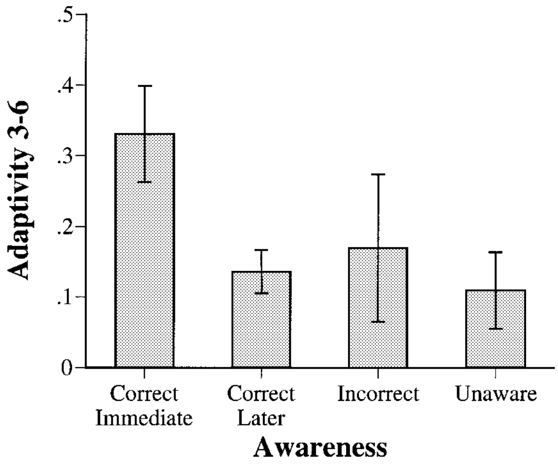

Figure 6. Mean strategy adaptivity (proportion overshoot in Block 3 minus proportion overshoot in Block 6) and *SE* for each explicit awareness group in Experiment 2.

using a summary of each participant's results to measure working memory capacity, we compared each participant's results on a memory task with the output of a computational model. The model posits that working memory capacity is a special kind of attentional activation directed at memory items related to the current goal, hence making these items more accessible relative to other items in memory. Within the model, there is a very precise quantitative relationship between the attentional/working memory parameter (W) and accuracy of recall, and this precise relationship depends on the precise timing details of the experiment (i.e., interstimulus delays, study–test lags, etc.). Thus, one can easily estimate the W parameter for a given person using his or her accuracy data in a well-controlled situation. Estimating working memory capacity in this fashion has been found to produce highly accurate and cross-situationally predictive estimates of working memory (Daily, Lovett, & Reder, in press; Lovett, Reder, & Lebiere, 1999).

Method

Participants. Forty-four CMU undergraduates participated for course credit and were randomly assigned in equal numbers to one of two conditions (high load or low load, to be described below). Due to a computer problem, BST data were not collected from 2 of the participants in the low-load condition.

Procedure. The problems for the BST were taken from the predictive condition of Experiment 2 (i.e., 60 problems that were all biased in looks, the looks were 73% predictive, and base rates shifted from 70% overshoot to 30% overshoot across blocks).

The same awareness-debriefing questionnaire was used as that in Experiment 2. In addition to this awareness questionnaire, the participants were asked to rate the frequencies with which they used four common strategies. Although the strategies were explained in more detail to the participants, the four strategies, briefly named, were the following: undershoot, overshoot, hill-climbing (i.e., select the stick that looks like it will bring you closest to the goal length), and select what worked on the previous trial. The participants rated their use of each strategy on a 5-point Likert scale, with the five points labeled: *always*, *usually*, *fairly often*, *rarely*, and *never*. The participants' responses were then recoded as a number from 1(*always*) to 5(*never*). The goal of asking the participants about their strategies was to examine whether the strategies that they used were systematically different as a function of working memory capacity or load condition.

All participants were given a secondary task to perform while solving the BST problems. This task involved processing and responding to auditorally presented items and, depending on the condition, maintaining certain items in memory. The items were a list of *a*s and *b*s presented at the rate of one every 3 sec. The participants in the low-load condition were asked to press the "z" key whenever an *a* was heard and the "x" key whenever a *b* was heard. The participants in the high-load condition were asked to press the "z" key whenever the current letter was the same as the previous letter and the "x" key whenever the current letter was not the same as the previous letter. Thus, the participants in the low-load condition did not need to remember what the auditory stimulus was after a response was made, whereas the participants in the high-load condition always had to remember what the previous auditory stimulus was.

As soon as a letter was presented auditorally, a red circle appeared at the bottom of the computer screen. The red circle indicated that a key response was required. If the participant made a correct response, the circle disappeared. If the participant made an incorrect response, the circle remained and the participant could make another response. The keypresses were made with the left hand; the right hand controlled the mouse, which was used for solving the BST problems.

The participants were given instructions and practice on the secondary task before beginning instructions on the BST. They practiced the secondary task until they made 20 consecutively correct responses or until 120 trials had passed. During the BST problem trials, the participants were given a self-timed rest period every 10 BST problems. During the rest period, the participants were also presented with statistics regarding their performance on both tasks: proportion of BST problems solved in five or fewer moves, proportion of secondary task trials answered correctly, and proportion of secondary task trials answered on time (i.e., before the next stimulus occurred). If less than 90% of responses (to the secondary task) were on time or if less than 80% of them were correct, the participants were encouraged to try harder.

The working memory task was also administered on the computer. This task was a modified digit span (MODS) task developed by Lovett et al. (1999) that is a variant of one developed by Oakhill and her colleagues (Yuill, Oakhill, & Parkin, 1989). In general, modified span tasks require participants to perform some other activity concurrently with the test of memory span (Daneman & Carpenter, 1980; Turner & Engle, 1989), and, as such, they tend to prevent participants from using different strategies that obscure differences in working memory capacity. In the MODS task, participants must read a sequence of letters and digits aloud while maintaining the digits in memory for later recall. Trials vary in the number of digits that must be recalled (from 3 to 6), and performance is measured as a proportion of trials at each set size that the participant recalls perfectly (i.e., correct digits in the exact order of presentation). These performance data are then used in conjunction with the computational model of the MODS task in order to estimate the best-fitting working memory capacity parameter for that individual (see Appendix B). The MODS task and the BST were administered in counterbalanced order across participants.

Results

Manipulation check. As an indication that the high-load condition was more difficult, the participants made

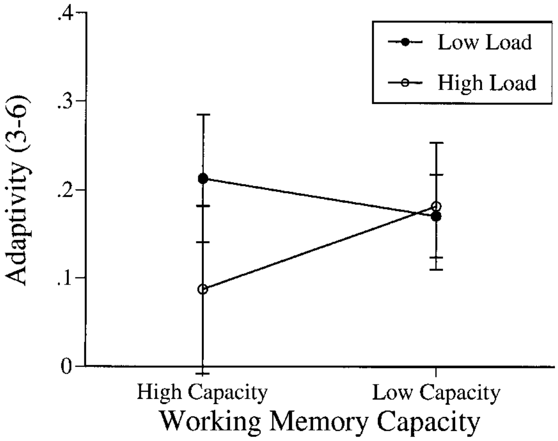

Figure 7. Mean strategy adaptivity (proportion overshoot in Block 3 minus proportion overshoot in Block 6) and *SE* as function of working memory capacity and condition in Experiment 3.

significantly more errors on the secondary task in the high-load condition than in the low-load condition (mean error rates of .26 and .11, respectively) [$F(1,40) = 7.2$, $MS_e = 0.028$, $p < .01$].

Overall adaptivity. To examine the effects of individual differences in working memory capacity, the participants were divided into high- and low-capacity groups using a median split on the *W* estimates obtained from the computational modeling fit, as described earlier.[2] A 2 (high capacity/low capacity) $\times$ 2 (high load/low load) ANOVA on adaptivity (mean overshoot use in Block 3 minus mean overshoot use in Block 6) found no effect of working memory capacity, no effect of condition, and no interaction (Fs < 1). As Figure 7 reveals, the only hint of an effect is that high-capacity participants in the high-load condition were less adaptive than were the low-capacity participants.

One might argue that the range of working memory capacity is restricted because all the participants were undergraduates at CMU. Yet, there have been many successful studies of individual differences in working memory using this same participant pool (e.g., Just & Carpenter, 1992; Lovett et al., 1999; Shah & Carpenter, 1995). Alternatively, one might argue that the large clumping of participants in a middle range of capacity dilutes the analyses. However, even if an extreme groups design is used in the analyses, throwing out the middle third of the working memory capacity participants, the effect of working memory capacity remains nonsignificant [$F(1,28) < 1$], with mean adaptivity levels of .11 and .12 for high- and low-capacity participants, respectively. Moreover, the correlation between adaptivity and working memory capacity is $r = -.02$.

Were the high- and low-capacity participants differentially allocating their resources? Examining performance on the secondary task, the high-capacity participants had neither higher accuracies (mean accuracies of .82 for both groups) nor faster latencies (mean median latencies of 0.83 and 0.84 sec for the high- and low-capacity groups, respectively) than the low-capacity participants (Fs < 1). Thus, both groups of participants appeared to devote equal resources to the secondary task. However, examining performance on the secondary task only in the high-load condition alone, the high-capacity participants had a weak trend toward higher accuracies (.81 vs. .69) [$F(1,16) = 1.99$, $MS_e = 0.03$, $p < .2$] and faster latencies (0.91 vs. 1.01 sec) [$F(1,16) = 1.51$, $MS_e = 0.03$, $p < .25$]. One could interpret this as simply a consequence of their higher capacities. Alternatively, one might interpret this trend as a tendency for the high-capacity participants to devote more resources to the secondary task in the high-load condition (and thus fewer resources to the primary task). Yet, if we remove any accuracy and latency differences by restricting our analyses to the participants with mean accuracies higher than .8, the high-capacity participants were still just as adaptive (if anything less adaptive) as the low-capacity participants in the high-load condition (mean adaptivities of .05 and .12; $F < 1$). Thus, it appears that the lack of differences in adaptivity is not likely to be attributable to differential allocation of resources.

Ideally, one would have wanted to have identical performance in the secondary task. Unfortunately, the high-load condition had lower performance levels on the secondary task. Thus, one might argue that the lack of differences in the BST across conditions was due to these performance differences (i.e., the participants adjusted their secondary task performance to maintain equal primary task performance levels). Subjectively, this explanation seems unlikely, since the high-load task was a much harder task than the low-load task, even with these compensations. The participants often reported forgetting how they had started a problem in the high-load condition, suggesting that it had effectively targeted working memory. Yet, to examine statistically whether the differences in secondary task performance could explain the lack of condition effects, an analysis of covariance was conducted predicting strategy adaptivity in the BST, with condition as a factor and secondary task accuracy as a covariate. Although there is a marginal predictiveness of secondary task accuracy [$F(1,37) = 1.97$, $p < .2$], there continues to be no suggestion of an effect of condition [$F(1,37) < 1$]. The same result was obtained when secondary task latencies were used as a covariate. Thus, the lack of differences in adaptivity across conditions appears not to be attributable to differences in secondary task performance.

Another possibility is that the high-capacity participants were using more demanding strategies to solve the building sticks problems than were the low-capacity participants. Similarly, the participants in the high-load condition may have switched to using less demanding strategies for the building sticks task. Thus, they may have been no more likely to notice the base-rate shifts, not because working memory capacity is irrelevant to strategy adaptivity but because of differential strategy use. To address this issue, the participants' were asked about their use of four different strategies: overshoot, undershoot,

Table 2
Number of Participants With Each Level of Awareness of the Base-Rate Changes Within Each Condition of Experiment 3

Condition	Correct Immediate	Correct Later	Incorrect	Unaware	Total
Low load	3	3	6	10	22
High load	2	5	3	10	20
Total	5	8	9	20	42

hill-climbing, and select what worked on the previous trial. Always selecting either overshoot or undershoot is the less demanding strategy, whereas hill-climbing requires spatial comparisons between goal and target, and selecting what worked previously requires a memory retrieval. However, there were no differences between either the high-capacity or low-capacity participants or the high-load or the low-load conditions in terms of the use of these four strategies (χ^2s < 1). Thus, it appears unlikely that the lack of differences among participant types or groups was caused by changes in strategy use.

Awareness and adaptivity. In contrast to Experiment 2, the great majority of the participants in Experiment 3 were unaware of the base-rate changes. Table 2 presents the number of participants with each level of awareness. Only 12% of the participants fell into the correct immediate group. This rate is considerably lower than the 28% categorized correct immediate in Experiment 2 [$\chi^2(1) = 3.6, p < .06$], which used essentially the same BST problems. Thus, the secondary task kept many of the participants from being able to attend to the base-rate changes.

Because so few participants were in the correct immediate group, there was little power in the analysis of the relationship between awareness and adaptivity. Therefore, it is not surprising that the effect of awareness on adaptivity (difference in overshoot use between Blocks 3 and 6) was not significant [$F(3,28) = 1.4$, $MS_e = 0.058$, $p < .3$] (see Figure 8). As in Experiment 2, the incorrect participants were likely to have been a mixture of correct immediate and unaware participants, and this may explain their high performance. Also as in Experiment 2, the correct later participants had adaptivity levels more similar in adaptivity to the unaware participants than to the correct immediate participants.

Discussion

Experiment 3 provided further evidence that working memory may not play a large role in adaptivity. As in Experiment 1, individual differences in working memory capacity did not correlate with individual differences in adaptivity. Moreover, the working memory load manipulation did not produce changes in strategy adaptivity either.

The adaptivity and awareness results were consistent with those of Experiment 2, although not as strong. The demands of the secondary task seemed to reduce awareness of the base-rate change. These same demands may also have reduced the ability of the aware participants to use the explicit awareness to influence their behavior (mean adaptivity levels of .33 and .22 for correct immediate participants in Experiments 2 and 3, respectively). By contrast, for the unaware participants, there were no differences in adaptivity (mean adaptivity levels of .10 and .09 in Experiments 2 and 3, respectively). Thus, it appears that, for unaware participants, the presence of a secondary task does not affect their more implicit strategy adaptivity.

One possible explanation of the reduction in explicit awareness in Experiment 3 is that working memory capacity is important for gaining explicit awareness. However, the addition of a secondary task does more than impact working memory resources; it also manipulates attentional focus and the amount of spare processing time one has. These other factors are plausible candidates for the source of the reduction in awareness. Moreover, the working memory capacity explanation is inconsistent with the lack of differences in awareness levels between the two conditions in Experiment 3.

GENERAL DISCUSSION

Across the three experiments, a consistent set of results emerged. First, strategy adaptivity does not appear to be directly related to working memory capacity (either individual differences in working memory capacity or through dual-task manipulations of available working memory). Second, explicit awareness of changes in the base rates of success appears to be related to larger strategy adaptivity to those changes. However, it is important to note that even the unaware do adapt to changes in base rates of success.

Both sets of findings come with a cautionary note. Although approached from a variety of methods, the results regarding memory capacity are null results and thus must be treated carefully. For example, we have not ruled out the possibility that base rates of success are stored in working memory but require so little resources that even

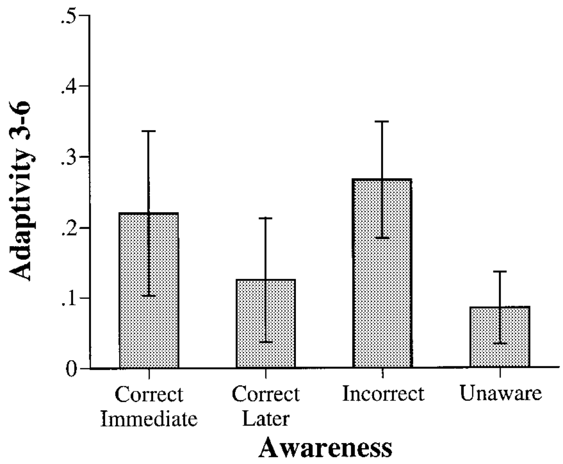

Figure 8. Mean strategy adaptivity (proportion overshoot in Block 3 minus proportion overshoot in Block 6) and *SE* for each explicit awareness group in Experiment 3.

low working memory span individuals in high memory load conditions still have enough working memory to store the base-rate information or that there were other tradeoffs (as suggested in Figure 7) that we were unable to identify statistically.

With respect to the role of awareness, it is important to note that all of our evidence is correlational and involves self-report data. It remains possible that increases in adaptivity produce increases in awareness level—that participants were simply reflecting on their own behavior and making attributions rather than noticing changes in the environment per se. Future research will have to study the effects of manipulating awareness on strategy adaptivity.

Similar to our findings, the implicit learning literature indicates a partial role for awareness in learning. For example, sequence learning can occur without explicit awareness of what is being learned (Nissen & Bullemer, 1987), but it does require that some attention is given to the items to be learned (Hartman, Knopman, & Nissen, 1989). Clark and Squire (1998, 1999) found that participants can be conditioned to noncontiguous eye-blink puffs and tones if and only if the participants are aware of the relationship between the tones and the puffs (although see LaBar & Disterhoft, 1998, for an alternative view).

Our findings suggest the following account of strategy choice. With enough time and thought, some participants (possibly the participants with greater inductive reasoning ability) become aware of base-rate changes. This awareness causes them to undergo a stronger shift in strategy preference, similar to Reder's (1988) findings that, with explicit instructions, participants can change their strategy preferences. By contrast, when participants are unaware of the base-rate changes, the strategy preferences do change, in a weaker more implicit fashion. This implicit strategy adaptivity is unaffected by differences in working memory capacity or by dividing attention with a secondary task.

Different aspects of this account are consistent with each of the previous accounts of strategy selection. The role of explicit awareness is consistent with the metacognitive accounts of cognitive control (Koriat, 1993; Metcalfe, 1994; Nelson & Narens, 1990). However, strategy adaptivity even in the absence of explicit awareness is consistent with the implicit accounts of strategy selection (Anderson & Lebiere, 1998; Lovett & Anderson, 1996).

Our present findings appear to be inconsistent with the claims that we have made previously that metacognitive control should be entirely implicit (Reder & Schunn, 1996). In that paper, we argued that strategy selection was governed by implicit memory, whereas here we have evidence that suggests a role of explicit awareness. However, the implicit control account is not entirely inconsistent with the present findings. First, people do adapt, even when unaware. In other words, we were partially correct. Second, in the present study, we focused only on strategy adaptivity to changes in base rates of success. People must also adapt their strategy use to other factors. For example, the looks of a problem, called an *intrinsic factor* (Reder, 1987), also influences strategy choice. For example, in our Experiments 2 and 3, the participants were influenced by whether the problem looked like it could be solved more quickly using undershoot versus overshoot strategies. In the BST domain, participants come with the prior expectations that the looks of the problem will predict which strategy will be successful. In other domains, however, participants must learn which features are predictive. It is differences in this learning that we call *differences in intrinsic adaptivity*. Schunn and Reder (1998) found that individual differences in intrinsic adaptivity appeared to be attributable entirely to differences in task expertise and speed–accuracy tradeoffs, whereas individual differences in extrinsic adaptivity were not attributable to those two factors. Thus, intrinsic adaptivity may be purely implicit, whereas extrinsic adaptivity is typically implicit but may be circumvented by awareness.

Our findings on the role of explicit awareness may have methodological consequences for both experimental and computational work on strategy adaptivity. It appears that aggregate performance data will always contain a mixture of aware and unaware participants, whose performance differs substantially. By modeling the aggregate, researchers may develop models of strategy selection that reflect the performance of no subgroup (Estes, 1956; Maddox, 1999; Siegler, 1987). Even worse, consider the case of a manipulation that has its effect primarily on whether participants are aware of a base-rate change rather than directly on the amount by which they adapt (e.g., the addition of a secondary task). A model that attempts to capture the effect of such a manipulation without taking into account the mediating awareness variable may be inaccurate.

The apparent lack of a relationship between working memory and explicit awareness in our findings is surprising. Initially, one might have expected either that both working memory and explicit awareness play a role in strategy selection or that neither play a role. However, our evidence suggests that explicit awareness, but not working memory, appears to play a role. Our findings raise the possibility that explicit awareness can act on information stored outside of working memory, a controversial claim that deserves additional research.

REFERENCES

Anderson, J. R., & Lebiere, C. (1998). *Atomic components of thought*. Mahwah, NJ: Erlbaum.

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.

Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: The role of awareness. *Science*, **280**, 77-81.

Clark, R. E., & Squire, L. R. (1999). Human eyeblink classical conditioning: Effects of manipulating awareness of the stimulus contingencies. *Psychological Science*, **10**, 14-18.

Daily, L. Z., Lovett, M. C., & Reder, L. M. (in press). Modeling individual differences in working memory performance: A source activation account. *Cognitive Science*.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory. *Journal of Verbal Learning & Verbal Behavior*, **19**, 450-466.
Davidson, J. E., Deuser, R., & Sternberg, R. J. (1994). The role of metacognition in problem solving. In J. Metcalfe & A. P. Shimamura (Ed.), *Metacognition: Knowing about knowing* (pp. 207-226). Cambridge, MA: MIT Press.
Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (2nd ed.). Cambridge, MA: MIT Press.
Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, **53**, 134-140.
Hartman, M., Knopman, D. S., & Nissen, M. J. (1989). Implicit learning of new verbal associations. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 1070-1082.
Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, **99**, 122-149.
Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, **100**, 609-639.
Kyllonen, P. C. (1993). Aptitude testing inspired by information processing: A test of the four-sources model. *Journal of General Psychology*, **120**, 375-405.
Kyllonen, P. C. (1994). Cognitive abilities testing: An agenda for the 1990s. In M. G. Rumsey, C. B. Walker, & J. H. Harris (Eds.), *Personnel selection and classification* (pp. 103-129). Hillsdale, NJ: Erlbaum.
Kyllonen, P. C. (1995). CAM: A theoretical framework for cognitive abilities measurement. In D. Detterman (Ed.), *Current topics in human intelligence: Volume IV. Theories of intelligence* (pp. 307-359). Norwood, NJ: Ablex.
LaBar, K. S., & Disterhoft, J. F. (1998). Conditioning, awareness, and the hippocampus. *Hippocampus*, **8**, 620-626.
Lemaire, P., & Reder, L. (1999). What affects strategy selection in arithmetic? The example of parity and five effects on product verification. *Memory & Cognition*, **27**, 364-382.
Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology*, **31**, 168-217.
Lovett, M. C., Daily, L. Z., & Reder, L. M. (2000). A source activation theory of working memory: Cross-task prediction of performance in ACT-R. *Cognitive Systems Research*, **1**, 99-118.
Lovett, M. C., Reder, L. M., & Lebiere, C. (1999). Modeling working memory in a unified architecture: The ACT-R perspective. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 135-182). Cambridge: Cambridge University Press.
Lovett, M. C., & Schunn, C. D. (1999). Task representations, strategy variability and base-rate neglect. *Journal of Experimental Psychology: General*, **128**, 107-130.
Luchins, A. S., & Luchins, E. H. (1950). New experimental attempts at preventing mechanization in problem solving. *Journal of General Psychology*, **42**, 279-297.
Maddox, W. T. (1999). On the dangers of averaging across observers when comparing decision bound models and generalized context models of categorization. *Perception & Psychophysics*, **61**, 354-374.
Metcalfe, J. (1994). A computational modeling approach to novelty monitoring, metacognition, and frontal lobe dysfunction. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 137-156). Cambridge, MA: MIT Press.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In A. C. Graesser & G. H. Bower (Eds.), *The psychology of learning and motivation: Inferences and text comprehension* (Vol. 25, pp. 125-173). New York: Academic Press.
Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Ed.), *Metacognition: Knowing about knowing* (pp. 1-25). Cambridge, MA: MIT Press.
Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, **19**, 1-32.
Raven, J. C., Court, J. H., & Raven, J. (1977). *Standard progressive matrices*. London: H. K. Lewis.
Reder, L. M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. *Psychological Review*, **89**, 250-280.
Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology*, **19**, 90-137.
Reder, L. M. (1988). Strategic control of retrieval strategies. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 227-259). San Diego, CA: Academic Press.
Reder, L. M., & Schunn, C. D. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. M. Reder (Ed.), *Implicit memory and metacognition* (pp. 45-78). Mahwah, NJ: Erlbaum.
Reder, L. M., & Schunn, C. D. (1999). Bringing together the psychometric and strategy worlds: Predicting adaptivity in a dynamic task. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance. Interaction of theory and application* (pp. 315-342). Cambridge, MA: MIT Press.
Schunn, C. D., & Reder, L. M. (1998). Strategy adaptivity and individual differences. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 38, pp. 115-154). New York: Academic Press.
Schunn, C. D., & Reder, L. M. (in press). Another source of individual differences: Strategy adaptivity to changing rates of success. *Journal of Experimental Psychology: General*.
Shah, P., & Carpenter, P. A. (1995). Conceptual limitations in comprehending line graphs. *Journal of Experimental Psychology: General*, **124**, 43-61.
Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, **116**, 250-264.
Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In G. Halford & T. Simon (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). New York: Academic Press.
Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229-293). Hillsdale, NJ: Erlbaum.
Spehn, M. K., & Reder, L. M. (2000). The unconscious feeling of knowing: A commentary on Koriat's paper. *Consciousness & Cognition*, **9**, 187-192.
Thorndike, E. L. (1913). *Educational psychology: The psychology of learning* (Vol. 2). New York: Columbia University, Teachers College.
Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, **28**, 127-154.
Yuill, N., Oakhill, J., & Parkin, A. (1989). Working memory, comprehension ability, and the resolution of text anomaly. *British Journal of Psychology*, **80**, 351-361.

NOTES

1. For this goodness-of-fit test, the .75–1.0 bin was pooled into the .50–.75 bin because the expected frequency was less than 5 for the higher bin.

2. The same results are found when working memory is treated as a continuous variable.

(*Continued on next page*)

APPENDIX A
Individual Difference Tests Used in Experiment 1

In the spatial working memory test, participants must process and remember a sequence of three 3×3 square matrices, each of which has only one square (of the nine possible) colored in. On a given trial, the participants are shown three matrices and must first decide whether the equation Matrix 1 + Matrix 2 = Matrix 3 is true or false (e.g., upper-left filled-in plus center filled-in equals a matrix with upper-left and center filled-in). After several such matrix equation trials are answered, participants are prompted to recall the matrices (each sum of Matrix 1 + Matrix 2) from that group of trials. A 3×3 matrix with numbers (1–9) in each square appears. Participants must type in the numbers one at a time corresponding to each of the squares shaded in the memorized matrices (e.g., 1 and 5 for a matrix with upper-left and center filled-in). To receive full credit, these numbers must be in the sequence in which the matrices were originally presented. The test is divided into two sets, each made up of groups of items containing 3, 4, and 5 memorized matrices and solved equations. The primary measure is the mean percentage of matrices recalled with all items in the correct order.

The three inductive spatial tests were as follows. In the figure sets test, participants are presented with three sets of figures. Two of the sets will be related according to various themes. Participants must determine which set is the odd set. Some of the various patterns include figures formed with straight lines as opposed to curved lines, internal shading versus no shading, and so on. There are 10 items that must be solved within a 5-min period. In the figure series test, participants are shown a series of shapes at the top of the screen and must choose the next shape occurring in the series from three numbered alternatives. As a simplified example, if the series was: “/ * // ___,” the answer would be “**.” There are 10 problems that must be solved within a 5-min period. In the figure matrices test, participants are shown a 3×3 matrix in which a figure is contained in all but one of the cells, similar to the Raven’s progressive matrices (Raven, Court, & Raven, 1977). There are patterns or rules that apply across the rows and down the columns of each matrix from which participants must induce what figure belongs in the empty cell. For example, if the three rows were “> >> >>>, | || |||, < << ___,” then the answer would be “<<<.” The matrix and eight alternative responses are shown on the screen simultaneously. Some of the rules and patterns used involve gradual shading of figures, successive additions or deletions to figures, rotation of figures, and so on. There are 9 problems that must be solved within a 10-min period.

APPENDIX B
The Procedure Used to Estimate Working Memory Capacity Parameter, *W*

The first steps in estimating the *W* parameter were completed before this experiment was conducted. These first steps include developing a computational model of the MODS task and establishing values for the other parameters in the model. Developing the model consists of specifying a set of processes for performing the MODS task that are designed to mimic those that participants use. These processes are specified in the ACT-R framework (Anderson & Lebiere, 1998). When the model is run through the simulated experiment (with the exact same timing details as the participants experienced), these processes are executed by a general set of mechanisms (provided by ACT-R) that enable the model to produce simulated actions (e.g., reading a letter, outputting a recalled digit) with particular probabilities and latencies. These simulated data (e.g., probability of digit recall) are then analyzed in a fashion similar to the participants’ data. Past research has shown that the model’s predictions offer a good fit to participants’ data, at both the aggregate level and the individual subject level (Daily et al., in press; Lovett, Daily, & Reder, 2000; Lovett et al., 1999).

Because our Experiment 3 used the same procedure and timing as this past work, the model and the global parameter values used previously could be used here without modification. We then apply the model to the present data set and estimate a best-fitting value of the *W* parameter for each participant’s individual data. This was accomplished by varying *W* from 0.7 to 1.3 (in increments of 0.05) and generating the model’s performance profile under each value of *W*. Here, a profile consisted of four numbers: the percentages of trials recalled perfectly for set sizes 3, 4, 5, and 6. These profiles were then matched to the corresponding data from participants in the present experiment. The *W* value that produced the smallest sum of squared errors between the model profile and each participant’s data was defined to be that participant’s estimated *W* value.

In simplest terms, the *W* value for each participant provides an overall measure of his/her performance on the MODS task. This measure is not equivalent to a simple average, however. *W* is a theoretically motivated measure of attentional capacity that relates to performance in a nonlinear way. Moreover, the *W* value estimated for a particular participant’s data takes into account the shape of his/her profile (i.e., the amount of decrease in performance for higher set sizes). Other work has shown that people with high versus low working memory capacity exhibit different performance profiles as working memory demands increase (Just & Carpenter, 1992).

(Manuscript received October 26, 1999;
revision accepted for publication June 6, 2000.)