Journal of Experimental Psychology: General
2001, Vol. 130, No. 4, 764–778

Copyright 2001 by the American Psychological Association, Inc.
0096-3445/01/$5.00 DOI: 10.1037//0096-3445.130.4.764

Task Preparation and Task Repetition: Two-Component Model of Task Switching

Myeong-Ho Sohn and John R. Anderson
Carnegie Mellon University

The switch cost (the disadvantage of performing a new task vs. a repeated task) has been attributed to lack of preparation for the switched task or priming of the repeated task. These sources were examined by manipulating foreknowledge of task transition (repeat or switch), response-to-stimulus interval (RSI), and practice level. Regardless of foreknowledge, the cost decreased with RSI and practice. The reduction was greater with foreknowledge than with no foreknowledge, and the amount of switch cost did not depend on foreknowledge. These results suggest that the switch cost with foreknowledge may consist of both inadequate preparation and repetition benefit but the switch cost with no foreknowledge may reflect repetition benefit only. An ACT–R (adaptive control of thought–rational) model was proposed, accommodating both preparation and priming effect with 2 independent processes: conflict resolution among productions and decay of chunk activation.

The course of human cognition is determined both by executive control and automatic control. Executive control is endogenous, goal-directed, intentional, and voluntary, reflecting the current goal. It has been referred to as a central executive (Baddeley, 1986), a supervisory attentional system (Norman & Shallice, 1986; Shallice, 1994), executive function (Logan, 1985), or controlled processing (Schneider & Shiffrin, 1977). An example of executive control is the foreknowledge effect. When a task goal is specified in advance, the task can benefit significantly from the endogenous preparation on the basis of foreknowledge even if the stimuli to be processed are not yet available (Carlson & Lundy, 1992; Sohn & Carlson, 1998). Automatic control, in contrast, is driven by an external stimulus or event that is strongly associated with a certain sequence of actions, regardless of the current goal (Norman & Shallice, 1986; Schneider & Shiffrin, 1977). An example is the priming effect: Processing the same or similar concept facilitates subsequent processing of another concept, regardless of whether it was intended (Neely, 1977). Another example is utilization behavior of frontal lobe patients who show capture-like automatic control. These patients do not seem to be able to suppress action sequences characteristically associated with everyday objects such as matches, scissors, and so forth, even though performing such actions is not appropriate and against their desire (Lhermitte, 1983).

The purpose of the current study is to investigate how executive and automatic control mechanisms have their effects when people rapidly switch from one task to another. More specifically, we are interested in the time courses of task preparation and task repetition effects. Task preparation can be achieved by giving people foreknowledge that they will perform a specific task; its effects reflect endogenous executive control. Task repetition can be achieved by having people perform the same task; its effects without foreknowledge of repetition reflect exogenous automatic control. We propose that preparation and repetition are independent and they tap different mechanisms: Repetition benefit occurs because the first performance of the task results in activation boost that makes the repeated task performance more efficient. In contrast, preparation has its effects through executive mechanisms that regulate whether a certain kind of information needs to be accessed. The effect of foreknowledge may be built up gradually as more time is permitted, whereas activation on the basis of priming may decrease as a function of time (Neely, 1977).

At the end of this article, we describe a realization of this proposal in the adaptive control of thought–rational theory (ACT–R; Anderson, 1993; Anderson & Lebiere, 1998). At a coarse temporal grain size, information processing in ACT–R can be characterized as symbolic, but at a finer temporal grain size, there are important subsymbolic processes supporting the symbolic level. At a symbolic level, ACT–R assumes that information processing involves a sequence of production rule firings. Each production rule involves retrieving some declarative information, called chunks, to transform the current goal state. Although processing is serial at the symbolic level, it is parallel at the subsymbolic level. At the subsymbolic level, multiple production rules compete to determine the course of information processing. We will propose that task preparation has its effects on this competition, such that, when prepared, a participant can skip retrieving information to determine the next task. Also at the subsymbolic level, retrieval of declarative information involves parallel activation processes. The speed of retrieval of a piece of information will depend on its level of activation. The activation of a declarative

---

Myeong-Ho Sohn and John R. Anderson, Department of Psychology, Carnegie Mellon University.

This work was supported by National Science Foundation Grant BCS-9975220. Parts of this research were presented at the 40th Annual Meeting of the Psychonomic Society, Los Angeles, November 1999. We thank Rich Carlson and Marsha Lovett for their helpful comments on drafts of this article. The source code of the model reported here is available by following the "Published ACT-R Model" link from the adaptive control of thought–rational (ACT–R) homepage (http://act.psy.cmu.edu).

Correspondence concerning this article should be addressed to Myeong-Ho Sohn, BH345E, Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213. Electronic mail may be sent to mhsohn@andrew.cmu.edu.

chunk reflects the recency and frequency with which it has been processed. For example, when a certain task has been performed recently, the declarative chunks necessary to perform the task may be still highly activated. If so, it would take less effort to repeat the retrieval of these chunks when the task is repeated. However, if a new task has to be performed, a new set of declarative chunks has to be retrieved, requiring more effort to do that.

Task Switching

The goal of the current study is to show that both task preparation and repetition are necessary for the most efficient task performance. Although human cognition seems to be perfectly capable of parallel processing in some situations, serial organization at the behavioral level provides an opportunity for more flexible and more efficient use of cognitive capacity. The serial organization inevitably requires frequent switching from one task to another and selection among multiple stimulus–response (SR) mappings available in a given situation. Because switching and selection involve such activities as planning at the strategic level, monitoring current cognitive processes, and allocating cognitive resources, cognitive control seems to be essential for task switching with multiple SR mapping rules. For this reason, the task-switching paradigm has been widely used to examine the control processes in human cognition.

The defining feature of this paradigm is that participants rapidly repeat the same task or alternate between different tasks. Figure 1 shows a schematic of the paradigm we used in this study. On every trial, a participant receives two stimuli, and each stimulus consists of one letter and one digit. Each stimulus requires a separate response, and a response-to-stimulus interval (RSI) separates the first response and the second stimulus. We refer to responding to the first stimulus as Task 1 and responding to the second stimulus as Task 2. Figure 2 describes the structure of the task. The color of the stimulus designates the relevant task that the participant has to perform with the stimulus. For example, if the stimulus is green, the task is to decide whether the letter is a consonant or a vowel (letter task). If the stimulus is red, the task is to decide whether the digit is an even number or an odd number (digit task). When Task 1 and Task 2 involve the same task, the trial is a task repetition. When they involve different tasks, the trial is a task switch.

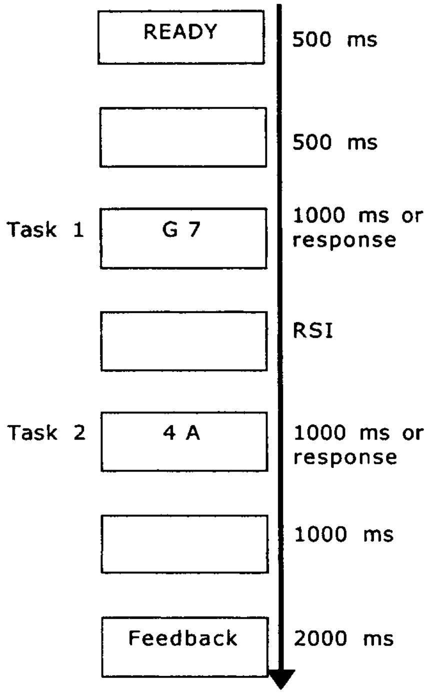

Figure 1. A schematic of a trial in Experiments 1 and 2. RSI = response-to-stimulus interval.

A consistent finding across different task-switching studies is that latencies are longer for task switches than for task repetitions (Allport, Styles, & Hsieh, 1994; Gopher, Armony, & Greenshpan, 2000; Jersild, 1927; Meiran, 1996; Rogers & Monsell, 1995; Sohn & Carlson, 2000; Spector & Biederman, 1976). This switch cost is remarkably robust even in a situation where a participant can supposedly prepare for a task switch, as long as the stimulus can afford two tasks as in the case of the current study. In many studies, there were two task sets and the task transition—whether to switch or to repeat—was predictable. Intuitively, if task switching is completely under the control of executive mechanisms, the cost should disappear when there is a sufficiently long RSI to prepare for the switched task. The switch cost was indeed reduced with RSI, indicating that at least some portion of the switch cost is due to lack of proper preparation for a task switch. However, Rogers and Monsell (1995) found that the cost reached an asymptotic level after 600-ms RSI. The cost was obtained even when the RSI was about 3,500 ms (Sohn, Ursu, Anderson, Stenger, & Carter, 2000). This residual switch cost implies some fundamental constraints on executive control when an individual rapidly switches from one task to another.

As a source of the residual switch-cost deficit, two independent but not mutually exclusive ideas have been proposed. One is inadequate preparation for a task switch compared with a task repetition. The idea is that the switch cost in general reflects the extent of readiness to perform the upcoming task. The reduction of the switch cost indicates that a person becomes prepared better with more time (Rogers & Monsell, 1995). Therefore, this view implies that the amount of switch cost is related to the extent of executive control over the to-be-performed task. This argument is based on the result that reduction of the switch cost was obtained only with predictable RSIs but not with random RSIs (Rogers & Monsell, 1995, Experiments 2 and 3). If the switch cost reflects some sort of automatic interference from the previous task, this interference should decrease as a function of time, whether or not the RSI is predictable. Instead, Rogers and Monsell's result suggests that the switch cost has something to do with knowing when the next stimulus will be presented, and therefore with preparation of the upcoming task. They suggest that the preparation for a switched task can be complete only on the arrival of the stimulus for the task, whereas the preparation for a repeated task has already been complete during the initial performance. This stimulus-cued completion hypothesis of task-set reconfiguration implies that the difference in the preparation process is responsible for the residual switch cost.

In accordance with the above idea, analyses of the cumulative distribution of response times of task switch and task repetition

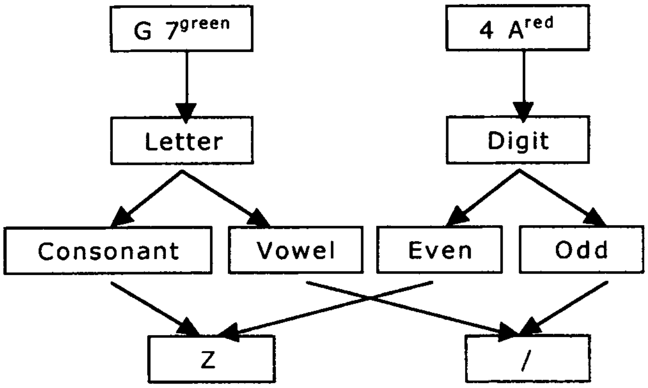

Figure 2. Color-to-task translation and category of stimulus-to-response mapping. The superscription indicates the color in which the stimulus was presented.

showed that with 1,500-ms RSI permitted for preparation, fast switch latencies were almost identical to fast repetition latencies (De Jong, 2000; De Jong, Berendsen, & Cools, 1999). This result suggests that, although preparation for a task switch can be as perfect as a task repetition, this does not happen all the time. In contrast, repetition latencies did not differ regardless of RSI, suggesting that preparation for repetition might be always nearly perfect. De Jong argued that switch and repetition trials differ in the demand for preparation and that preparation sometimes fails on switch trials, which is consistent with the position that inadequate preparation is the source of the switch cost.

Another proposed source for the residual switch cost is persisting activation from a previous task. Allport et al. (1994; Allport & Wylie, 1999, 2000; Wylie & Allport, 2000) suggested that the reduction of a switch cost during RSI reflects decaying activation of a previous task set. According to their task-set inertia (TSI) hypothesis, this activation never disappears completely until there is a new external stimulus that designates a new task, eventually resulting in the residual switch cost. The residual activation may facilitate performance of the repeated task but may interfere with performance of a switched task, resulting in the switch cost. Although the TSI hypothesis does not deny that the switch cost may also reflect the lack of executive control, it certainly denies that the reduction of switch cost directly reflects the increasing readiness for the upcoming task (Allport & Wylie, 2000). The idea is that the switch cost is due to the processes related to the previous task, not the upcoming task (Allport & Wylie, 1999; Wylie & Allport, 2000). For example, the difficulty of previous task had effects on the switch cost but not the difficulty of the upcoming task (Allport & Wylie, 1999; however, see Mayr & Kliegl, 2000). According to the TSI hypothesis, the task switch cost indicates the extent of interference from the previous task: The more difficult to perform the task, the more difficult to switch to another task from it.

Two-Component Hypothesis of the Switch Cost

One purpose of the current study is to show that although both inadequate preparation and repetition benefit may be causes of the general switch cost, the residual switch cost is due only to repetition benefit. This two-component hypothesis would indicate that the repetition and preparation are separate processes. In this sense, our study is a continuation of those studies that have focused on either preparation or repetition benefit (Allport et al., 1994; Rogers & Monsell, 1995) and shares its interest with those studies that have investigated multiple components of the switch cost (Meiran, 2000; Meiran, Chorev, & Sapir, 2000). Meiran and his colleagues, for example, have shown that the switch cost was affected by two different timing parameters. One is the time for repetition benefit from a previous task to dissipate, and the other is the time for the upcoming task to be prepared. However, these two intervals are not completely independent, because the repetition benefit effect can still dissipate while the upcoming task is prepared. We manipulated other variables orthogonal to the timing so that differences between repetition and preparation can be detected when the time permitted for each process is the same.

We measured the switch cost when participants perform two tasks, with the second task switched (task switch) or repeated (task repetition) from the first task. This task-transition manipulation will allow us to examine the effect of repetition over switch, the switch cost. In addition, we manipulated foreknowledge in a task-switching situation. With foreknowledge, participants knew whether they would perform a task repetition or a task switch. With no foreknowledge, they performed both task transitions randomly intermixed. The expected difference in reaction times between foreknowledge and no-foreknowledge conditions is the foreknowledge effect. The foreknowledge manipulation will allow us to examine the effect of preparation on the basis of foreknowledge. The motivation for the foreknowledge manipulation is that with foreknowledge, both preparation and task repetition contribute to the switch cost, whereas only task repetition contributes to the switch cost with no foreknowledge (Sohn & Carlson, 2000). This is because with no foreknowledge, a preparation for a task repetition cannot be different from the preparation for a task switch.

To test the independence of the preparation and repetition effects, we examined whether these effects vary over different lengths of time interval and also over different amounts of practice. The RSI interval between Task 1 response and Task 2 stimulus was 200 ms, 600 ms or 1,500 ms in Experiment 1, and 1,000 ms, 3,000 ms, and 5,000 ms in Experiment 2. The extent of preparation will increase with RSI, whereas the benefit from repetition will decrease with RSI (Neely, 1977). Therefore, the foreknowledge effect will increase with RSI regardless of task transition, whereas the switch cost will decrease with RSI regardless of foreknowledge.

The two-component hypothesis we propose here suggests that the reduction of switch cost with RSI is due to different reasons in each foreknowledge condition. For a repetition with no foreknowledge, the priming benefit will decrease with RSI. For a switch with no foreknowledge, there is no source of benefit. Therefore, with no foreknowledge, the switch cost will decrease because the repetition trial gets slower with RSI. With foreknowledge, the readiness will increase with RSI for a switch trial. Also, with foreknowledge, there will be a decreasing priming benefit as well as an increasing preparation for the repetition trial. Therefore, with foreknowledge, the switch cost will decrease because the switch trial becomes faster with RSI. This rather complex pattern of results cannot be predicted by simply assuming a single component model of the switch cost. For example, if the switch cost reduction reflects only decreasing activation from the previous task, this should be realized by the same pattern of reduction regardless of foreknowledge. Also, if the switch-cost reduction reflects only increasing preparation for a task switch, this would not be sufficient to capture the pattern of results in the no-foreknowledge condition. It is because neither repetition nor switch can be systematically prepared when there is no foreknowledge.

Regarding the effect of practice, we have an assumption that both preparation and repetition effects become smaller as participants become more fluent with tasks. It is because participants will achieve general speed up with practice regardless of experimental conditions, and each effect will be proportional to the baseline performance. Note that there are two sources of switch-cost reduction in the foreknowledge condition (preparation and repetition), whereas there is only one source of switch cost reduction in the no-foreknowledge condition (repetition). Therefore, the two-component hypothesis suggests that the practice should have greater impact on the performance with foreknowledge than should the performance with no foreknowledge: The reduction of the switch cost with practice will be greater with foreknowledge than with no foreknowledge.

Furthermore, the current design allows us to examine the source of the residual switch cost. As mentioned earlier, the switch cost in the no-foreknowledge condition should reflect only the repetition benefit, whereas the switch cost in the foreknowledge condition should reflect both preparation and repetition benefit. We expected that the switch cost would decrease with RSI and practice, eventually reducing to the residual switch cost especially with relatively long RSIs and substantial practice. What would be the source of the residual switch cost? One idea is that the residual switch cost reflects the persisting activation from the previous trial (Allport et al., 1994; Sohn & Carlson, 2000). Because this activation does not depend on preparation, the amount of the residual switch cost should not be different depending on the foreknowledge condition. Alternatively, the residual switch cost could reflect the inadequate preparation for a task switch compared with a task repetition (De Jong, 2000; Rogers & Monsell, 1995), and the activation benefit of a task repetition may disappear eventually with time. If this is true, not only the switch cost should be greater with foreknowledge than with no foreknowledge, but the residual switch cost should also disappear in the no-foreknowledge condition. This is because in the no-foreknowledge condition, no preparation is possible either for a repetition or for a switch.

Experiment 1

Table 1 illustrates the foreknowledge manipulations in experiments reported here. In the foreknowledge condition, participants performed a block of trials consisting only of task switches or task repetitions. In contrast, in the no-foreknowledge condition, they performed a block of trials involving both task switches and task repetitions randomly mixed within a block. Before each block started, participants received explicit instructions on the specific type of block they would perform. Also, RSI between Task 1 and Task 2 varied. For a specific block of trials, only one RSI was used (200 ms, 600 ms, or 1,500 ms), and participants were informed about this. To examine the effect of practice, three sessions were administered, with each session separated by a day.

Table 1
An Illustration of the Foreknowledge Manipulation in Experiments 1 and 2

Condition	Type of block	Task 1-Task 2
Foreknowledge	Repetition	Letter-Letter Digit-Digit
	Switch	Digit-Letter Letter-Digit
No foreknowledge	Mixed	Letter-Letter Letter-Digit Digit-Letter Digit-Digit

In Experiment 1, we are testing the following predictions. First, the switch cost will decrease as a function of RSI regardless of foreknowledge, but the pattern of reduction will be different depending on foreknowledge. With foreknowledge, the reduction of the switch cost should be achieved by decreasing latency of task switch trials. This is because the preparation for a task switch should increase with RSI. In contrast, with no foreknowledge, the reduction of the switch cost should be achieved by increasing latency from task-repetition trials. This is because the repetition benefit decreases with RSI. Second, the switch cost will be reduced with practice, but this practice effect will be more evident with foreknowledge than with no foreknowledge, because practice can affect both task repetition and task preparation. In both foreknowledge and no-foreknowledge conditions, practice can strengthen the knowledge that is being retrieved and so reduce the task-repetition effect. Also, in the foreknowledge condition, participants can learn to prepare more with practice and so reduce the preparation component of the switch cost. Third, if the residual switch cost reflects the repetition benefit, then the switch cost with foreknowledge will eventually reduce to the level of the switch cost with no foreknowledge. However, if the residual switch cost reflects lack of preparation, the switch cost will completely disappear with no foreknowledge but will remain with foreknowledge.

Method

Task and equipment. Stimuli were generated using MEL software (Micro Experimental Lab system, Psychology Software Tools, Pittsburgh, PA), and the timing was controlled by an IBM compatible PC. A stimulus consisted of one letter and one digit. The letter was one of four consonants (G, K, M, and R) or four vowels (A, E, I, and U). The digit was one of four even numbers (2, 4, 6, and 8) or four odd numbers (3, 5, 7, and 9). The left–right position of letter and digit was random. A participant's task was to make a decision about whether the letter was a consonant or a vowel (letter task) or whether the digit was an even number or an odd number (digit task). The task identity was indicated by the color of the stimulus. For example, if the stimulus was green (red), the task was the letter (digit) task.

Procedure and design. A participant initiated a block by pressing the space bar. As shown in Figure 1, every trial began with the READY signal at the center of the computer screen. The READY signal stayed for 500 ms and was replaced by a blank screen for 500 ms. Then, the Task 1 stimulus was presented on the screen for 1,000 ms, and the screen remained blank until the participant pressed a response key. For responses, the *Z* and the "/" keys were used. After an RSI, the stimulus for Task 2 appeared for 1,000 ms, and after that duration, the screen remained blank until the participant pressed a response key. After the Task 2 response, the screen remained blank for another 1,000 ms. Then, feedback about accuracy and latency for both tasks was presented for 2,000 ms, which was replaced by the READY signal for the next trial. The "even" and "consonant" responses were assigned to the same response key, and the "odd" and "vowel" responses were assigned to another response key. The mapping between responses and response keys was counterbalanced across participants.

In the first session, participants were given a practice block of 32 trials that consisted of all possible task transitions randomly intermixed. Then, they received 12 experimental blocks of 32 trials each. The organization of trials in a given block was different depending on the foreknowledge condition and the transition condition. As in Table 1, in the foreknowledge-repetition block, Task 1 and Task 2 were always the same. In the foreknowledge-switch block, Task 1 and Task 2 were always different. Therefore in these blocks, although Task 1 was not predictable, information about Task 2 became available as soon as the Task 1 stimulus was presented. In the no-foreknowledge blocks, task repetitions and task switches were randomly mixed. Before each block began, participants were instructed on the type of the block. The block order was constrained so that foreknowledge and no-foreknowledge blocks alternated. For example, foreknowledge repetition, no foreknowledge, foreknowledge switch, no foreknowledge, foreknowledge repetition, and so on. A sequence of 4 blocks was associated with the same RSI, and there were 12 blocks in each session. The RSI order was randomized for each participant. In the second and the third sessions, there were 12 main blocks without the practice block. Each session was separated by a day from each other. All independent variables were manipulated for all participants, resulting in a 3 (session) × 3 (200 ms, 600 ms, and 1,500 ms RSIs) × 2 (foreknowledge and no foreknowledge) × 2 (switch and repetition) within-subjects design.

Participants. A total of 32 college students, graduate students, and staff members at Carnegie-Mellon University participated in return for monetary reward, consisting of base pay and bonus proportionate to the speed on correct trials. The reward ranged from $27 to $33 for three sessions.

Results and Discussion

In this section, we report only latency results for simplicity of presentation. Accuracy results are reported in the Appendix. Accuracy data did not vary in a way that would compromise the interpretation of the latency results. In all of the experiments reported here, only correct trials were included in analyses of latency, and a trial was counted as correct when responses for Task 1 and Task 2 were both correct. Average accuracy and Task 2 latency data were negatively correlated across the various conditions defined by the combinations of practice sessions, RSI, foreknowledge, and transition ($R = -.91$, $p < .0001$), indicating that conditions with poor performance by one measure had poor performance by the other. The main analyses were four-way analyses of variance (ANOVAs) with practice (Sessions 1, 2, and 3), RSI (200 ms, 600 ms, and 1,500 ms), foreknowledge condition (foreknowledge and no foreknowledge), and task transition (repetition and switch) as factors for Task 1 and Task 2 latency. Although Task 1 latency is not our main interest, we analyzed these data to examine whether there is any carryover effect from the previous trial. For Task 1 latency, therefore, the task transition (or carryover) was whether the current Task 1 was the same as the Task 2 of the previous trial. For Task 1 latency, the first trial of a block was eliminated from the analyses. For Task 2 latency, the task transition was whether the current Task 2 was the same as the Task 1 of the same trial.

Task 1 latency. The ANOVA model applied to Task 1 is 3 (session) × 3 (RSI) × 2 (foreknowledge) × 2 (carryover). Mean Task 1 latency was 855 ms. The main effect of practice was significant, $F(2, 62) = 90.00$, $p < .0001$, $MSE = 112{,}709.20$. Task 1 latency was slowest in Session 1 (1,038 ms), intermediate in Session 2 (802 ms), and fastest in Session 3 (726 ms), Newman-Keuls, $p < .01$. The main effect of carryover was significant, $F(2, 62) = 35.01$, $p < .0001$, $MSE = 17{,}209.75$. Task 1 latency was faster when the current Task 1 was the same as the previous Task 2 (832 ms) than when it was different (878 ms). This carryover effect seems to depend on foreknowledge as reflected in the significant interaction between foreknowledge and carryover, $F(2, 62) = 16.40$, $p < .01$, $MSE = 3{,}796.38$. However, t tests showed that the carryover effects with foreknowledge (60 ms) and with no foreknowledge (31 ms) were both significant, $t(31)$, $p < .0001$, and $t(31)$, $p < .01$, for foreknowledge and no foreknowledge, respectively. No other main effects or interactions were significant, $p > .10$. Although it is not clear why the carryover effect is greater with foreknowledge condition, Task 1 latency results suggest that the carryover has its effect regardless of foreknowledge.

Task 2 latency. Mean Task 2 latency was 706 ms. Table 2 provides a complete list of the ANOVA for reference to the F values, degrees of freedoms, significance levels, and *MSE*s. Figure 3 shows the mean Task 2 latency as a function of RSI, foreknowledge, and task transition in each practice session. In Figure 4, Figure 3 is replotted in terms of the switch cost, which is the difference between task repetition and task switch. The main effects of practice session, foreknowledge, and task transition were significant. Participants responded to Task 2 slowest in Session 1 (825 ms), intermediate in Session 2 (670 ms), and fastest in Session 3 (623 ms), Newman-Keuls, $p < .01$. Participants were faster with foreknowledge (664 ms) than with no foreknowledge (748 ms). Participants were faster with task repetitions (627 ms) than with task switches (785 ms).

The interaction between RSI and foreknowledge was significant. As the Figure 3 shows, foreknowledge effect increased as RSI increased from 200 ms (75 ms) and 600 ms (78 ms) to 1,500 ms (100 ms), Newman-Keuls, $p < .01$. Also, interaction between

Table 2
Analysis of Variance Results of Experiment 1

Source	*F*	*df*	*MSE*
Practice Session (S)	131.25***	2, 62	32,665.39
RSI (R)	1.12	2, 62	25,533.68
Foreknowledge (F)	107.53***	1, 31	18,927.93
Transition (T)	135.91***	1, 31	52,601.23
S × R	0.52	4, 124	26,269.70
S × F	6.48**	2, 62	6,075.22
S × T	22.85***	2, 62	12,254.61
R × F	7.06**	2, 62	2,574.61
R × T	31.03***	2, 62	5,813.04
F × T	18.91**	1, 31	6,841.93
S × R × F	0.32	4,124	2,503.32
S × R × T	0.95	4,124	6,134.83
S × F × T	6.74**	2, 62	3,453.41
R × F × T	0.65	2, 62	3,121.95
S × R × F × T	0.76	4,124	2,813.25

Note. RSI = response-to-stimulus interval.
** $p < .01$. *** $p < .0001$.

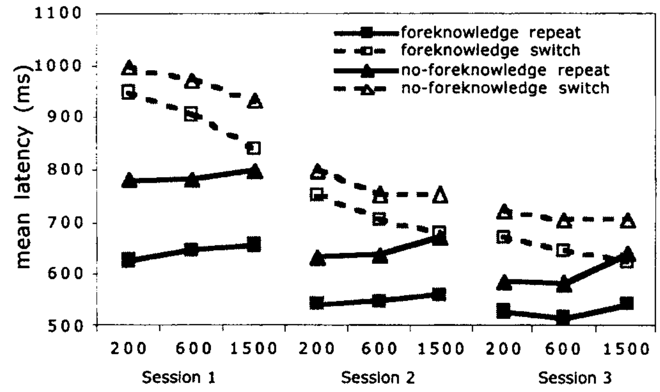

Figure 3. Task 2 latency as a function of response-to-stimulus interval in each foreknowledge and transition condition from each session in Experiment 1.

RSI and transition was significant. The amount of the switch cost (i.e., the effect of task transition) was greatest with 200-ms RSI (198 ms), intermediate with 600-ms RSI (162 ms), and smallest with 1,500-ms RSI (112 ms), Newman-Keuls, $p < .01$.

The three-way interaction involving practice session, foreknowledge, and task transition was significant. To identify the locus of the three-way interaction, we examined whether the switch cost depended on foreknowledge in each session. As in Figure 4, the switch cost depended on foreknowledge in Sessions 1 and 2, but not in Session 3. In Session 1, the switch cost with foreknowledge (256 ms) was greater than the switch cost with no foreknowledge (180 ms), $t(31) = 3.66, p < .01$. Also in Session 2, the switch cost with foreknowledge (159 ms) was greater than the switch cost with no foreknowledge (121 ms), $t(31) = 4.09, p < .01$. However, in Session 3, the switch cost with foreknowledge (121 ms) was not significantly different from the switch cost with no foreknowledge (108 ms), $p > .09$.

In Experiment 1, three results are noteworthy. First, the foreknowledge effect increased with RSI, indicating that the likelihood of preparation in the foreknowledge condition increased as participants had more time. Second, the switch cost decreased with RSI, indicating that the activation of the previous task decayed over time. Third, the switch cost with foreknowledge was generally greater than with no foreknowledge, indicating that the switch cost with foreknowledge reflected more than automatic priming. The extra component of the switch cost with foreknowledge can be attributed to inappropriate preparation for a task switch compared with a task repetition. When the RSI is relatively short or the amount of practice is small, participants may likely be better prepared for a task repetition than for a task switch. However, as the more time is permitted for preparation or as they become more fluent, their preparation for a task switch will become more efficient, reducing the amount of switch cost due to inadequate preparation.

The two-component hypothesis of the switch cost was supported by significant three-way interaction involving practice session, foreknowledge, and transition: The switch cost reduction with practice was greater with foreknowledge than with no foreknowledge. This interaction indicates that with practice, participants become more effective at preparing in the switch condition with foreknowledge, and this condition no longer suffers a preparation disadvantage relative to the repetition with foreknowledge. Further supporting evidence for the two-component hypothesis would have been a three-way interaction involving RSI, foreknowledge, and transition. In the switch with foreknowledge condition, the switch cost should decrease with time, both because participants can become more prepared and the repetition benefit decreases. Therefore, there should be a greater decrease in the switch cost with RSI in the foreknowledge condition than in the no-foreknowledge condition, where there is only the repetition priming factor. However, although there was a hint (Figure 4), this interaction was not significant. The lack of three-way interaction was further tested in Experiment 2.

The current results imply that the residual switch cost may be due to the persisting activation from the previous trial. This is because the amount of the switch cost in the foreknowledge condition reduced to the same amount as the switch cost in the no-foreknowledge condition, as the RSI became longer and also the practice level increased. As Figure 4 shows, reduction of the switch cost was greater in the foreknowledge condition than in the no-foreknowledge condition. We hypothesize that this is because there are two decreasing components in the switch cost with the foreknowledge, whereas there is only one with no foreknowledge. However, the fact that the amount of the switch cost did not differ eventually suggests that the preparation component in the foreknowledge condition may be saturated as practice and RSI increase. Therefore, the remaining switch cost should be attributed to the repetition benefit on the basis of the persisting activation, consistent with Allport et al. (1994).

Experiment 2

Experiment 1 results indicate that the switch cost may have different components depending on foreknowledge. With foreknowledge, both inadequate preparation and repetition priming would contribute to the switch cost, whereas only repetition priming benefit would contribute to the switch cost with no foreknowledge. Therefore, all things being equal, practice should have greater impact on the switch cost with foreknowledge than on the switch cost with no foreknowledge, because the former has two sources to improve but the latter has only one. Supporting this, the

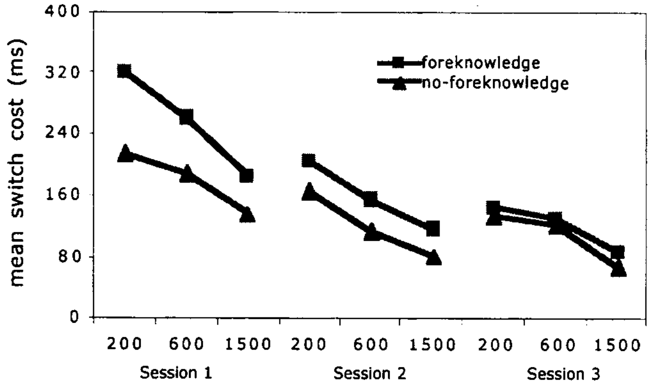

Figure 4. Switch cost as a function of response-to-stimulus interval in each foreknowledge condition from each session in Experiment 1.

three-way interaction involving practice, foreknowledge, and transition was significant in Experiment 1.

The purpose of Experiment 2 was to further generalize this two-component hypothesis in a different situation. Namely, we sampled different RSIs. Our motivation to select 200 ms, 600 ms, and 1,500 ms RSIs in Experiment 1 was on the basis of Rogers and Monsell's (1995) study in which the switch cost reached an asymptotic level with RSIs longer than 600 ms. However, on hindsight, the asymptotic RSI in our paradigm might have been different because of several differences between our paradigm and theirs. For example, they used a long sequence of tasks in which letter tasks and digit tasks were periodically repeated (e.g., . . .letter-letter-digit-digit-letter-letter. . .), and the identity of each task was predictable. In contrast, in our paradigm, there were two tasks in a sequence, and only the second task was predictable in the foreknowledge condition. The short sequence and the limited predictability may have required longer time to prepare for Task 2. In Experiment 2, we sampled an extremely long RSI (5,000 ms) along with shorter ones (1,000 ms and 3,000 ms), expecting to see the same effects as in Experiment 1 but even further reduction of the switch cost at the very long RSIs.

We also tried to replicate the practice effect in an experiment with a greatly reduced length. Instead of three sessions, we administered only one session. In each session of Experiment 1, one RSI was performed for four consecutive blocks and never repeated within the same session. Although RSIs were repeated across sessions and the order of RSIs was counterbalanced across participants, performing a chunk of blocks with the same RSI might have desensitized participants to the RSI manipulation. In Experiment 2, we reduced the number of trials that were associated with the same RSI within a block. Relatively speaking, Experiment 1 adopted massed practice of RSI, whereas Experiment 2 adopted distributed practice of RSI, so that the practice effect could be examined in a short period of practice.

Method

Task and equipment. Task and equipment were the same as in Experiment 1.

Procedure and design. The procedure was the same as in Experiment 1 with two changes. First, as noted earlier, different RSIs (1,000 ms, 3,000 ms, and 5,000 ms) were sampled. Second, as in Experiment 1, each block contained 32 trials and there were 12 blocks, but the organization of the 32 trials in each block was different. In Experiment 1, all 32 trials involved the same conditions, either repetition with foreknowledge, switch with foreknowledge, or no foreknowledge. In Experiment 2, the 32 trials in a block were subblocked so that each group of 8 consecutive trials would involve a different condition of foreknowledge and task transition. Therefore, participants received all combinations of foreknowledge and transition in each block. Every block was associated with different RSI, and the RSI order was randomized for each participant with the constraint that the same RSI would not repeat without other RSIs appearing between the repetitions. Each RSI was repeated four times so that we could compare performance from the first half of the experiment with performance from the second half. Resulting design was a 2 (first and second halves) × 3 (1,000-ms, 3,000-ms, and 5,000-ms RSIs) × 2 (foreknowledge and no foreknowledge) × 2 (switch and repetition) within-subjects design.

Participants. A total of 24 college students, graduate students, and staff members at Carnegie-Mellon University participated for a monetary reward of $10 base pay (or 1 hr course credit) plus a bonus depending on performance. The bonus ranged from $3 to $8.

Results and Discussion

Accuracy and Task 2 latency data for each combination of practice half, RSI, foreknowledge, and transition were negatively correlated ($R = -.83$, $p < .0001$). The main analyses were four-way ANOVAs with practice (first and second halves), RSI (1,000 ms, 3,000 ms, and 5,000 ms), foreknowledge condition (foreknowledge and no foreknowledge), and task transition (repetition and switch) as factors for Task 1 latency and Task 2 latency. As in Experiment 1, the task transition for Task 1 was whether the current Task 1 was the same as the Task 2 of the previous trial. For Task 2 latency, the task transition was whether the current Task 2 was the same as the Task 1 of the same trial.

Task 1 latency. Mean Task 1 latency was 1,041 ms. In Experiment 2, both Task 1 and Task 2 latencies were slower than in Experiment 1, possibly because of less practice. Task 1 latency was faster in the second half (873 ms) than in the first half (1,210 ms), $F(1, 23) = 137.28$, $p < .0001$, $MSE = 119{,}229$. Task 1 latency was faster with carryover (1,016 ms) than with no carryover (1,067 ms), $F(1, 23) = 13.94$, $p < .01$, $MSE = 26{,}434$. No other main effects or interactions were significant, $p > .09$. Task 1 latency results suggest that carryover effect was not different depending on foreknowledge conditions.

Task 2 latency. Mean Task 2 latency was 908 ms. Table 3 shows the complete list of ANOVA results of Task 2 latency from Experiment 2. Figure 5 shows the mean Task 2 latency as a function of RSI in each foreknowledge and task transition condition in each half. In Figure 6, Figure 5 is replotted in terms of the switch cost. Task 2 latency was faster in the second half (802 ms) than in the first half (1,013 ms). Task 2 latency was faster with foreknowledge (827 ms) than with no foreknowledge (988 ms). Task 2 latency was faster with repetition (859 ms) than with switch (957 ms).

The interaction between RSI and foreknowledge was significant. Foreknowledge benefit with 5,000-ms RSI (214 ms) was greater than with 3,000-ms RSI (154 ms) or 1,000-ms RSI (115 ms), Newman-Keuls, $p < .01$. The interaction between RSI and transition was significant, $F(2, 46) = 12.50$, $p < .0001$,

Table 3

Analysis of Variance Results of Experiment 2

Source	*F*	*df*	*MSE*
Practice half (H)	70.18***	1, 23	91,407
RSI (R)	2.02	2, 46	43,475
Foreknowledge (F)	81.58***	1, 23	45,706
Transition (T)	32.73***	1, 23	42,202
H × R	0.20	2, 46	43,325
H × F	13.13**	1, 23	11,170
H × T	4.95*	1, 23	17,362
R × F	16.65***	2, 46	7,209
R × T	12.50***	2, 62	14,902
F × T	7.52*	1, 23	14,902
H × R × F	2.04	2, 46	12,927
H × R × T	0.25	2, 46	10,205
H × F × T	10.79**	1, 23	6,183
R × F × T	1.19	2, 46	13,976
H × R × F × T	0.83	2, 46	14,249

Note. RSI = Response-to-stimulus interval.
*$p < .05$. **$p < .01$. ***$p < .0001$.

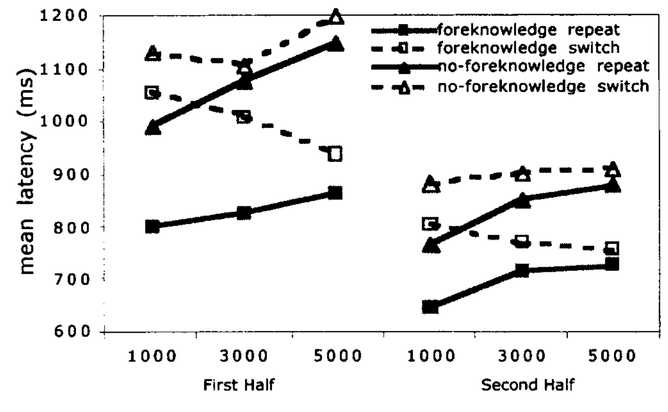

Figure 5. Task 2 latency as a function of response-to-stimulus interval in each foreknowledge and transition condition from each half in Experiment 2.

MSE = 15,090. The switch cost with 1,000-ms RSI (168 ms) was greater than with 3,000-ms RSI (80 ms) or 5,000-ms RSI (46 ms), Newman-Keuls, $p < .01$. As we suspected, the switch cost continued to decrease with the longer RSIs of this experiment.

The three-way interaction involving practice, foreknowledge, and transition was significant. In the first half, the interaction between foreknowledge and transition was significant, because of the fact that the switch cost was greater with foreknowledge (172 ms) than with no foreknowledge (73 ms), $t(23) = 3.435, p < .05$. However, the same two-way interaction was not significant in the second half, $p > .40$, indicating that the switch cost did not depend on foreknowledge in the second half.

Experiment 2 replicated Experiment 1 results very well with dramatically different RSIs. First, the foreknowledge effect increased with RSI. Second, the switch cost decreased with RSI. Third, the switch cost with foreknowledge was reduced with practice to a greater extent than the switch cost with no foreknowledge. Fourth, the switch cost in the foreknowledge condition reduced to the same amount as the switch cost with no foreknowledge. The results support the two-component hypothesis of the switch cost with foreknowledge. The preparation component may be more affected by practice and does not exert its effects anymore when the task is well practiced. This fact supports the idea that the residual switch cost may be due to the repetition benefit. However, once again, the three-way interaction involving RSI, foreknowledge, and transition was not significant, even though this effect was very suggestive in Figure 5. We will return to the issue of this three-way interaction in the following section after presenting an ACT–R model for our experiments.

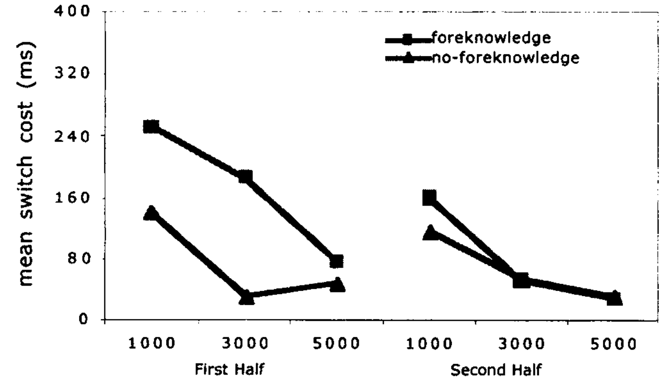

Figure 6. Switch cost as a function of response-to-stimulus interval in each foreknowledge condition from each half in Experiment 2.

Two-Component ACT–R Model of Task Switching

We developed an ACT–R model that simulates the current results. ACT–R is a cognitive theory that assumes two systems of long-term memory, declarative knowledge and procedural knowledge. The two systems of knowledge prove to be exactly what we need to model the two components of the switch cost. The declarative component will hold information such as the mapping of the color onto task or the mapping of categories (e.g., odd and even) onto responses. Repetition of such declarative components will provide the repetition priming benefit with or without foreknowledge. The procedural component will be responsible for setting the goal to do one task or another. It will be responsible for the preparation for the task switch during the RSI. The unit of declarative knowledge is a chunk, and the unit of procedural knowledge is a production. The declarative chunks critical to the current simulation are presented in Table 4. Productions consist of conditions (listed after *IF*) and actions (listed after *THEN*). The English version of productions of the task-switching model is presented in Table 5.

The model starts by encoding the color of the stimulus (Start-Task) and translating the color into the corresponding task (Encode-Task). The translation involves retrieval of one of two task-translation chunks (see Table 4). Once the task is identified, the model identifies the relevant symbol, letter or digit (Identify-Symbol). The identified letter (digit) symbol is categorized further as consonant or vowel (even or odd) by the production Judge-Symbol. The model then maps the category of the symbol to a response (Map-Response). This response mapping involves another chunk retrieval of a stimulus-to-response mapping rule (see Table 4). Once the response is identified, the model executes the response (Respond). This basic sequence of production firings applies to both Task 1 and Task 2, but our modeling effort is focused on Task 2 performance. One difference between Task 1 and Task 2 is that, if the participant is prepared, the Task-Prepared

Table 4
Declarative Chunks Critical to Task-Switching Model

Chunk type and chunk name	Chunks
Color-to-task translation	
Translation 1	Green indicates a letter task.
Translation 2	Red indicates a digit task.
Stimulus-to-response mapping	
Mapping 1	Consonant is indicated by the "z" key.
Mapping 2	Vowel is indicated by the "/" key.
Mapping 3	Even is indicated by the "z" key.
Mapping 4	Odd is indicated by the "/" key.
Task 1-to-Task 2 switch	
Switch 1	A letter task switches to a digit task.
Switch 2	A digit task switches to a letter task.

Table 5
The English Version of Task-Switching Productions

Production	Conditions and actions
Start-Task	IF goal is to perform a trial, and stimulus is available THEN encode color
Encode-Task	IF color is encoded, and the task is not known and translation chunk with the encoded color is found THEN set the task as the task value from the Translation chunk and encode stimulus (letter and digit)
Task-Prepared	IF the task is known and the stimulus is available THEN encode stimulus (letter and digit)
Identify-Symbol	IF stimulus is encoded (letter and digit) THEN identify the symbol (letter or digit) for the task
Judge-Symbol	IF the symbol is identified THEN categorize the symbol (e.g., consonant or vowel)
Map-Response	IF the symbol is categorized and category-to-response-mapping chunk is found THEN set the response as the response value from the Mapping chunk
Respond	IF response is known THEN respond, and ready for another
Prepare-Switch	IF transition is switch, and in the RSI, and task-inverse chunk is found THEN set the task as the next-task value from the Inverse chunk
Think	IF in the RSI THEN think

Note. RSI = response-to-stimulus interval.

production can apply instead of the Encode-Task production as the task is already known. This saves retrieval of the task-translation chunk.

Normally, no task-relevant productions apply during the RSI. In the foreknowledge-repeat condition, a participant is already prepared for Task 2 and in the no-foreknowledge condition, the participant cannot prepare. However, in the foreknowledge-switch condition, the participant can prepare for the upcoming task. To do this, the Prepare-Switch production must fire to determine what Task 2 will be. This requires that it retrieves a chunk encoding the other task (see Table 4). For the model to be prepared for Task 2, the Prepare-Switch production must apply before the Task 2 stimulus is presented. Once the stimulus is presented, the Start-Task production will fire to initiate Task 2 and either the Encode-Task or Task-Prepared production will fire depending on whether Task 2 is prepared.

The RSI will be filled with various ongoing processes (e.g., reflecting on Task 1 performance, considering other aspects of the experiment, or thinking about nonexperimental issues). To prepare for a switch trial with foreknowledge, it is necessary for the Prepare-Switch production to intrude itself into this ongoing processing. Given their potentially wide variety, we do not attempt to model these ongoing processes. Rather, we assume a production will fire on average every 200 ms and the simulation simply has a single generic Think production that fires every 200 ms as a stand-in for this ongoing stream of thoughts. Our assumption is that there are thought processes that are naturally evoked by the end of Task 1 and participants are inclined to follow these during the RSI. The Prepare-Switch production will intrude only with difficulty. The participant may find it much easier to prepare for Task 2 once the Task 2 stimulus is presented, which amounts to the assumption that exogenous control tends to dominate over endogenous control but that this dominance can be overcome. Note that in this respect our model is very similar to the theory of De Jong (2000), who holds that it is possible but difficult for a participant to prepare fully for a task switch.

The latency to perform a task is determined by the time for the productions to fire to execute the task. According to the ACT–R theory, each production takes a minimum of 50 ms to fire. The first production will take additional time because it awaits perception of the stimulus and the last will take longer because it executes a response. All of these are fixed times that do not vary across conditions. The variable time involves retrieval of information from declarative memory. The two critical productions in this regard are Encode-Task that retrieves the mapping between color and task and Map-Response that retrieves the mapping between stimulus category and response.

Priming, Conflict Resolution, and Practice

There are two ways that judgment time can be sped up in ACT–R. First, it is possible to prime the retrievals by recently performing them. Second, it is possible to prepare for the task in advance and fire the Task-Prepared production rather than the Encode-Task production. With respect to priming, a chunk-retrieval time depends on the activation of the chunk, and the chunk activation is a function of retrieval frequency and recency. The chunk activation equation (Anderson & Lebiere, 1998) is

$$A = \ln\left(\sum_j t_j^{-d}\right), \tag{1}$$

where t_j is the time elapsed since the jth time that chunk was accessed and d is the decay rate. All other things being equal, the more frequently or the more recently a chunk has been retrieved, the higher the activation. In particular, there will be a marked advantage of having retrieved the same chunk in the previous task as in a task repetition. The activation of a chunk in turn determines the time to retrieve the chunk according to the equation

$$\text{Time} = Fe^{-A}, \quad (2)$$

where F is a latency scale to be estimated. In the current model, the repetition benefit is produced by the activation level of the retrieved declarative chunk. By performing a certain task in Task 1, the activation level of the relevant chunks for the task becomes higher than others. When these chunks are retrieved again in Task 2, there is a benefit from this extra activation, resulting in faster reaction times. However, activation decreases with RSI, resulting in smaller repetition benefit with RSI.

Conflict resolution is the term that is used to describe the selection process that applies in ACT–R whenever multiple productions compete to apply. In ACT–R, the probability of a production winning the conflict resolution depends on its utility relative to the utilities of the other competing productions. Utility means how useful the production is in terms of achieving the current goal. Because of noise in the system, it is not always the production with the highest utility that is selected. The equation to calculate the probability of a production i with the utility U_i winning a conflict resolution among n applicable productions with utilities U_j is

$$\text{Probability of } i = e^{U_i} / \sum_j e^{U_j}, \quad (3)$$

where the summation is over the n alternatives.

To prepare in the foreknowledge switch condition, the Prepare-Switch production must win in the conflict resolution with the other ongoing productions. As the RSI increases, there is an increasing probability that the model will be prepared, because every 200 ms there is another conflict resolution between the Think production and the Prepare-Switch production. Note that the current model assumes rather a discrete, all-or-none type of preparation. For the model to be prepared, it is sufficient that the Prepare-Switch production wins only one of the many conflict resolutions. This is because, once the Prepare-Switch production wins, the content of the goal changes accordingly, so that the model knows which is the next task. Therefore, the RSI effect on the switch cost reduction can be simulated by an increasing probability of preparation with RSI.

This all-or-none preparation with probabilistic increase, consistent with De Jong (2000), is not the assumption shared by all the task-switching theories. Instead, there has been at least an implicit assumption that the level of preparation increases gradually (e.g., Rogers & Monsell, 1995). That is, the longer the RSI, the greater the extent of preparation. It is as if activation of the task is gradually built up to some extent with RSI. The gradual increase of preparation on the basis of activation build-up implies quite a different mechanism for the preparation than the discrete preparation on the basis of chunk retrieval (Mayr & Kliegl, 2000). However, it is not our intention to theoretically commit to either kind of preparation, because two kinds of preparation are functionally equivalent in terms of producing RSI effects in our experiments.

The practice effect was stimulated in two ways. First, the activation level of relevant chunks increased from session to session (Experiment 1) or from the first half to the second half (Experiment 2). This activation-level increase affected fluency in all four conditions. Second, the probability of preparation increased also from session to session or from the first half to the second half, which specifically affects the fluency in the switch with foreknowledge condition. This was done by increasing utility of the Prepare-Switch production.

Model Fitting

We fit this ACT–R model simultaneously to the data (60 data points) from both experiments. In fitting the model we estimated 18 parameters, presented in Table 6. Of these parameters, 6 reflected intercept time, which is the time for the stimulus to be detected, the productions to apply, and the response to be generated. Because of the different RSIs and slightly different experimental procedures, we estimated one intercept parameter for each RSI in each experiment. These intercepts in Table 6 show that participants were somewhat slower in Experiment 2 and that there

Table 6
Estimated Parameters for ACT–R Simulation

Experiment	Condition	Estimate
	Intercept	
Experiment 1	RSI = 200 ms	362 ms
	RSI = 600 ms	340 ms
	RSI = 1,500 ms	340 ms
Experiment 2	RSI = 1,000 ms	418 ms
	RSI = 3,000 ms	429 ms
	RSI = 5,000 ms	431 ms
	Probability	
Experiment 1	Session 1	.119
	Session 2	.128
	Session 3	.149
Experiment 2	First half	.080
	Second half	.123
	Base	
Experiment 1	Session 1	.386
	Session 2	.597
	Session 3	.705
Experiment 2	First half	.331
	Second half	.534
	Decay	
Experiments 1 and 2		1.25
	F	
Experiments 1 and 2		178 ms

Note. ACT–R = adaptive control of thought–rational; RSI = response-to-stimulus interval; F = latency factor.

was a tendency to have slightly faster times when the RSI was around 1 s.

Next we estimated the probability that the participant would fire the Prepare-Switch production rather than the Think production during the RSI in the switch condition with foreknowledge. In the actual simulation, this is controlled by the utilities, but the utilities are really just set to estimate the probability of the Prepare-Switch production firing in every 200-ms interval. We assumed that the participants might have increased their tendency to prepare with practice, and so we estimated separate parameters for each practice condition. Indeed, the probability did increase with practice (see Table 6). Note also that the probabilities for the two halves of Experiment 2 bracket the probabilities for the first session of Experiment 1.

The remaining seven parameters concerned the retrieval of declarative information about the color-to-task translation or the stimulus category-to-response mapping. This is controlled by the activation of the chunks, and there is a particularly strong effect of having accessed that knowledge during Task 1. The activation of a chunk can be closely approximated by the following equation modified from the Equation 1:

$$A = \ln(\text{Base} + t^{-d}), \tag{4}$$

where Base reflects the base for that session (Equation 1), and t^{-d} reflects how much the item has decayed since it was last accessed t s ago. Because there were four category-to-response mappings and only two color-to-task translations, we constrained the base levels for the color-to-task translations to be twice the base levels of the category-to-response mappings. Thus, we had five base parameters to estimate for the category-to-response mappings corresponding to the three sessions in Experiment 1 and two halves of Experiment 2. The base activations for the color-to-task translation chunks are just twice the value of the category-to-response mapping. These base levels are well behaved, with increasing values as practice increases and with the Experiment 2 values bracketing the Session 1 value for Experiment 1. Finally we estimated a decay value, d, of 1.25 and a latency factor, F, of 178 ms.

These parameters were estimated to minimize the chi-square deviation of the fit to the data defined as

$$\sum_{i} (\hat{X}_i - \bar{X}_i)^2/s_{\bar{x}}^2, \tag{5}$$

where $\hat{X}_i$ is the predicted mean, $\bar{X}_i$ is the observed mean, and $s_{\bar{x}}^2$ is the error of the mean estimated from the Participant × Condition interaction for that session (Experiment 1) or half (Experiment 2). This is a statistic that has a number of degrees of freedom (42) equal to the number of observations (60) minus the number of parameters (18). The actual chi-square measure of the deviations is 37.77, which is not significant. The simulation for reaction times from Experiments 1 and 2 are presented in Figures 7 and 8, and switch costs are presented in Figures 9 and 10. The overall correspondence with the data is quite good. The R^2 between theory and data is .973 for Session 1 of Experiment 1, .988 for Session 2 of Experiment 1, .978 for Session 3 of Experiment 1, .948 for the first half of Experiment 2, and .997 for the second half of Experiment 2.

Note that the model reproduces all of the significant effects in the experiments. It produces the foreknowledge effect because of preparation. It also produces the switch cost effect because of both repetition benefit and inadequate preparation for a switch with foreknowledge. It speeds up with sessions mainly because of the increased base activations of the chunks but also because it is more often prepared in the switch-with-foreknowledge condition. The foreknowledge effect increases with RSI because of greater probability of preparation, and the switch effect decreases because of decay of activation. It also reproduces the three-way interaction between session, foreknowledge, and transition because it is more often prepared in the switch-with-foreknowledge condition with practice, reducing the switch cost with foreknowledge.

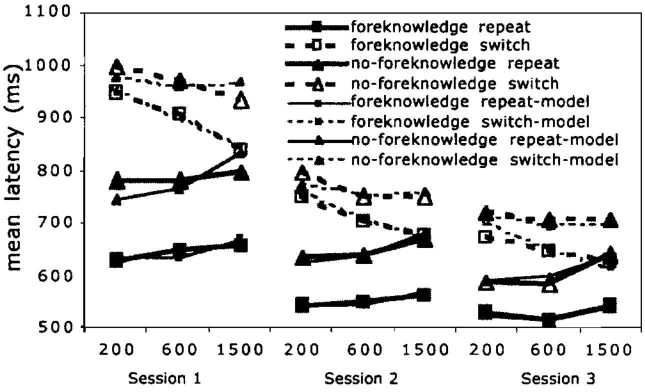

Figure 7. Simulation of Task 2 latency from Experiment 1. Thick lines represent data, and thin lines represent predictions.

In addition to all these effects, the model also predicts a three-way interaction between RSI, foreknowledge, and transition for the same reasons it predicts the other three-way interaction. With longer RSIs it should be more prepared, and this should reduce the switch costs with foreknowledge more than the switch cost with no foreknowledge. However, whereas suggestive, this three-way interaction was not significant in either experiment. Perhaps our selection of RSIs in each experiment was not comprehensive enough to reveal the three-way interaction involving foreknowledge, transition, and RSI. However, across the two experiments, we used a relatively larger variation in RSI from 200 ms to 5,000 ms. Figure 11a displays the switch costs from the Session 1 of Experiment 1 and the average switch costs of the first and the second halves of Experiment 2 plotted together as a function of RSI. This reflects data of comparable practice from each experiment. Combining the short RSIs from Experiment 1 with the long RSIs from Experiment 2 makes quite transparent ACT–R prediction of an interaction between RSI, foreknowledge, and transition. The data appear to support this interaction. Figure 11b plots the observed difference in the switch cost between foreknowledge and no foreknowledge as a function of the predicted difference by ACT–R. The correlation between observed and predicted was .83, which is significant at the .05 level for six points. Thus, over the two experiments, there is a significant tendency to display the predicted reduction in the switch cost difference between foreknowledge and no foreknowledge across different RSIs.

Regarding the residual switch cost, the current model implicitly assumes that the switch cost will keep decreasing and eventually disappear with sufficiently long time and practice, rather than reaching an asymptote. In fact, with a long enough RSI, both our

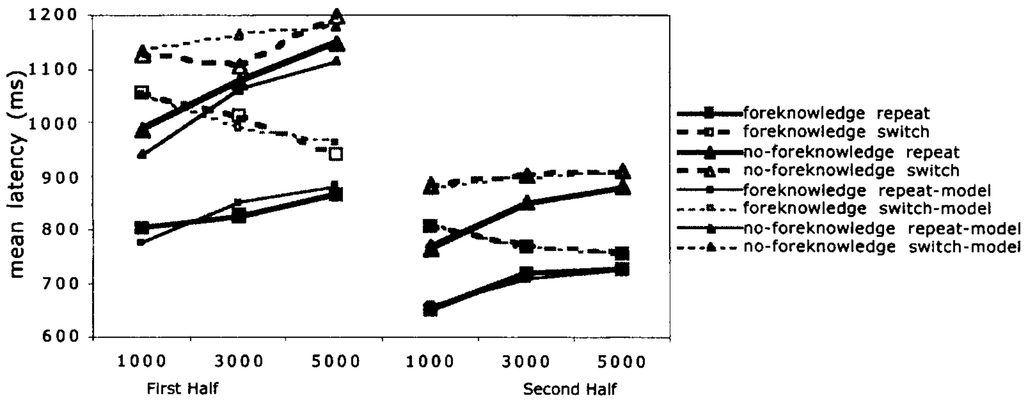

Figure 8. Simulation of Task 2 latency from Experiment 2. Thick lines represent data, and thin lines represent predictions.

data and the model suggest that the switch cost approaches zero, as indicated by Figure 11. However, the activation equations provided earlier, Equations 1 and 4, indicate that although the reduction of activation should be rapid during the early period of decay interval, it would take an infinite amount of time for the activation to truly go to zero. Perhaps, the residual switch cost may reflect this slowly decaying activation after some amount of time.

General Discussion

Three results are important in the current study. First, the foreknowledge effect increased with RSI. During RSI, participants seemed to prepare for Task 2 with foreknowledge, but not with no foreknowledge. The increase of the foreknowledge benefit with RSI supports the conclusion that foreknowledge allows voluntary preparation for Task 2, which is presumably controlled by an executive mechanism. Second, the task switch cost decreased with RSI, indicating that, whatever causes the switch cost, its effect decreases as more time is permitted between two tasks. Third, the foreknowledge effect on the switch cost depended on practice level. Early on in the practice, the switch cost was greater with foreknowledge than with no foreknowledge. This difference almost disappeared as participants practiced more.

Altogether, the above results suggest that preparation and repetition are both necessary conditions to efficiently prepare for a task, but how they affect the switch cost may depend on another factor, such as practice. The separate contribution of task preparation and task repetition was simulated in an ACT–R model. In this model, task-preparation effect was simulated by advance retrieval of a critical declarative chunk, and the repetition effect was simulated by the activation boost of a chunk that was recently retrieved and would be retrieved again. The advance retrieval reflects executive control in the sense that this is controlled by the task goal. The activation boost reflects automatic control in the sense that this has its effect regardless of intention and its effect involves interpretation of the stimulus. For example, if the upcoming task is the same as the previous one, the boosted activation will allow less effort to be spent on stimulus encoding. However, if the upcoming task is different from the previous one, more effort needs to be used to process the stimulus.

The current data and the model support the idea that the switch cost consists of two components, the lack of preparation for a task switch and the priming benefit of a task repetition. Both of these components decreased as RSI and practice level increased, as reflected in the reduction of switch cost. This is consistent with previous studies suggesting that the switch cost reflects both increasing preparation and decreasing interference (Meiran, 2000). To our knowledge, however, no studies have ever dissociated the time course of these components as we did. Moreover, the current study sheds a doubt on the ideas that advocate ill preparation as an important source of the residual switch cost. In our study, the switch cost persisted even with long RSIs and substantial practice, resulting in the residual switch cost. However, the amount of the residual switch cost did not depend on the foreknowledge. Consistent with previous studies (Dreisbach, Haider, & Kluwe, 2001; Sohn & Carlson, 2000), this equal amount of the residual switch cost seems to suggest that its source may be the persisting activation from the previous trial rather than the lack of preparation. This is because the residual switch cost in the no-foreknowledge condition cannot be attributed to the preparation difference but to the repetition benefit.

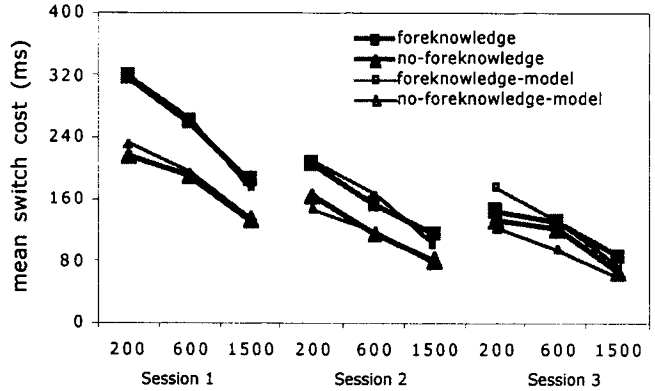

Figure 9. Simulation of switch cost from Experiment 1. Thick lines with filled markers represent data, and thin lines with open markers represent predictions.

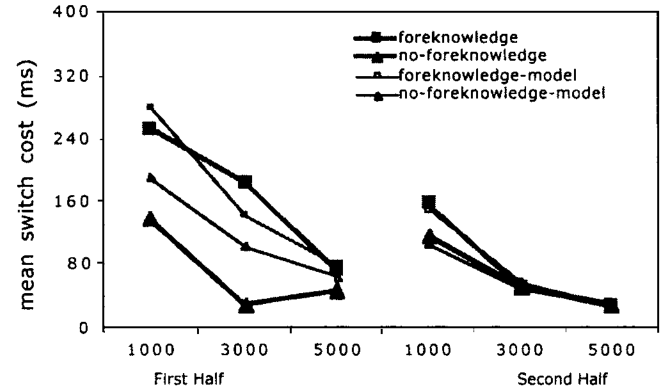

Figure 10. Simulation of switch cost from Experiment 2. Thick lines with filled markers represent data, and thin lines with open markers represent predictions.

A functional magnetic resonance imaging study also supports the assumption of our ACT–R model that the switch cost involves separate preparation and repetition components. In this study (Sohn et al. 2000), it was shown that task-preparation effect and task-repetition effect were associated with different brain areas. When the task foreknowledge was available, inferior prefrontal cortex (Brodmann Area [BA] 46/45) showed greater activation before the task was actually presented. Moreover, the higher the activation in this area, the faster the subsequent task performance, suggesting that this area is associated with task preparation. In contrast, when no foreknowledge was available, superior prefrontal cortex (BA 8) was activated more with a task switch than with a task repetition. The higher activation in this area was correlated with slower responses, suggesting that less effort may be required to perform the repeated task than to perform a new task. This dissociation supports the idea that executive control (e.g., task preparation) and automatic control (e.g., task repetition) are both crucially and yet independently involved in task performance.

What does it mean that task preparation and task repetition are both crucial for efficient task performance? Perhaps, the dual control system of executive and automatic mechanisms may maintain a balance between endogenous and exogenous impacts on selecting the course of actions. For example, foreknowledge may prepare a participant to focus on the task-relevant aspect of a stimulus so that the disturbance due to the task-irrelevant aspect of the stimulus can be minimized. This view is consistent with evidence from frontal lobe patients, whose executive control mechanisms are disrupted because of the brain damage. The utilization behavior mentioned earlier may be the consequence of the remaining influence of automatic control after the disruption of executive control. Also, studies with the Wisconsin Card Sorting Test showed that frontal lobe patients had difficulty dealing with task-irrelevant aspects of stimulus (e.g., Owen et al. 1993). These patients showed a greater difficulty sorting the cards when a new dimension became relevant, but the previously relevant dimension remained as a distractor dimension. Although speculative, the current results imply that these patients might be more successful in a slower paced task, because the automatic effect from the task-irrelevant dimension will decrease with time.

Throughout the experiment, foreknowledge resulted in a substantial latency benefit for both repetition and switch. This benefit was assumed to be due to preparation for an upcoming task. The result that the foreknowledge effect increased with RSI indicates that the likelihood of preparation gradually increases over time. Exactly what is prepared about the upcoming task? How specific can the preparation be in the task switching? The current model assumes a general kind of preparation. Basically, what our model does for preparation is to set itself for the upcoming task by changing the task value in its goal before the actual stimulus arrives. Although they are not a part of the current modeling effort, other more specific kinds of preparation are also possible. Rogers and Monsell (1995) characterized the preparation as engagement or reconfiguration of a cognitive system for a new task set. The task set specifies SR mapping rules, for example, “If a letter is a consonant, then press the ‘z’ key.” In our experiments, there were four SR mapping rules. With foreknowledge, the number of relevant rules decreases from four to two. As an alternative to retrieval of a task-translation chunk, participants might have rehearsed this subset of the SR mapping rules during preparation in the foreknowledge condition.

Exactly what is facilitated by repetition? One interesting aspect of repetition benefit in a task-switching paradigm is that repeating the same response from the previous task has different impact on

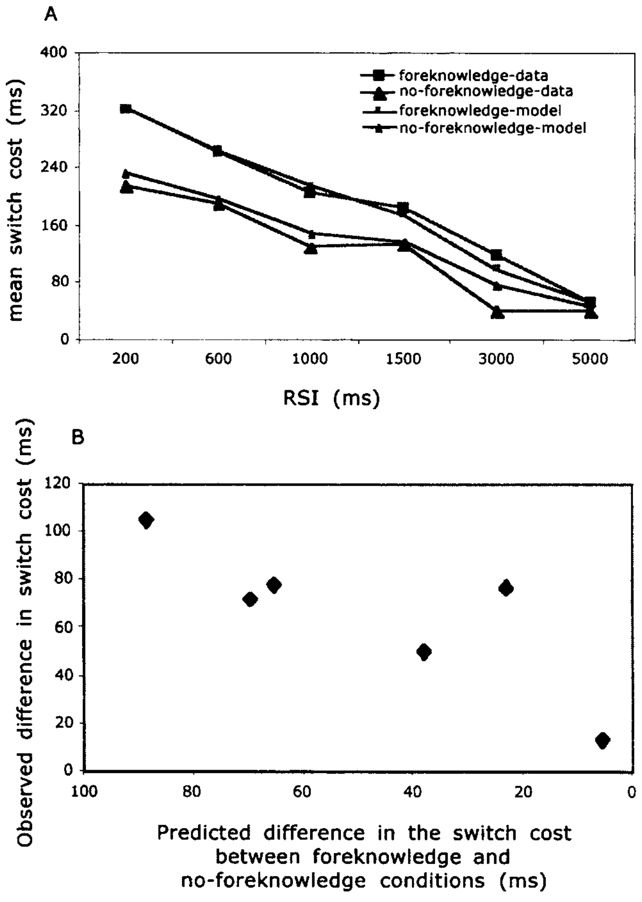

Figure 11. A: Switch cost from Experiment 1 (Session 1) and Experiment 2 (average of first and second halves) as a function of response-to-stimulus intervals (RSIs). B: Observed switch cost as a function of predicted switch cost by ACT–R (adaptive control of thought–rational).

a task repetition and a task switch. Other researchers (Mayr & Keele, 2000; Rogers & Monsell, 1995) have found that response repetition results in a benefit in the context of task repetition, but response repetition results in a cost in the context of task switch. Although we did not report it for simplicity of results section, we also conducted the same analysis and found similar results. In the current model, the repetition benefit comes from two sources. One is repeating the same color-to-task translation, and the other is repeating the same SR mapping rule. These two sources interact in the sense that the repetition of SR mapping rule results in a benefit only when the color-to-task translation is repeated. That is, the response repetition benefit is due to repeating the same SR mapping rules, but it is not due to repeating the same response per se. Because switch of task sets necessarily means switch of SR mapping rules, our model predicts neither benefit nor cost of response repetition with a task switch. This discrepancy of the model suggests modifications to elaborate the processes involved in response generation.

Sohn and Carlson (2000) found no interaction between task transition and foreknowledge. In our results, this was achieved with longer RSIs with longer practice. The implication of the same amount of the switch cost regardless of foreknowledge is that the source of the residual switch cost may be the persisting activation from the previous task set. Actually, the ACT–R model does not predict a fixed ordering of switch costs in the conditions of foreknowledge and no foreknowledge. The switch cost will reflect the retrieval of task translation and response mapping chunks. The effects associated with retrieval of response mapping will be the same regardless of preparation—participants will have a priming benefit when the same response is used in the repetition condition. However, the effects associated with the task translation vary as a function of condition. In the no-foreknowledge condition, there will be a difference in the priming of the task translation as a function of whether or not the task is repeated. In the foreknowledge condition, the task translation does not even have to be retrieved when the task is repeated. However, depending on preparation in the switch-with-foreknowledge condition, the participant may be able either to skip retrieval of the task translation or have to retrieve an unprimed task translation chunk. As the probability of preparation in this condition varies from 0 to 1, the size of the switch cost will vary from greater than the no-foreknowledge condition to less than the no-foreknowledge condition. In our data and model, the switch costs with and without foreknowledge were approximately equal in conditions of long RSI where preparation was relatively high (see Figure 11). The Sohn and Carlson experiments used relatively long RSIs, and this may be why they found no interaction between task transition and foreknowledge.

In conclusion, this research has found evidence that the switch cost reflects both effects of executive control and automatic control. Also, this study showed that the residual switch cost mainly reflects the repetition benefit. The executive control involves preparing for the task in the foreknowledge condition. According to the ACT–R model, the participant already starts prepared in the repeated condition but in the switch condition the participant has to divert normal processing during the RSI to prepare for the task. This executive switch is more likely to happen as RSI increases or as the participant gains more practice. The automatic control involves priming of the experimental information from the previous trial. Priming can affect both the availability of information about the color-to-task mapping and the stimulus category-to-response mapping. The repetition priming will decay during the RSI and become less important as the participant practices the mapping rules. Thus, the ACT–R model predicts a complex mix of executive and automatic effects. The rich pattern of data in Figures 3–10 conforms to these predictions.

References

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). Cambridge, MA: MIT Press.

Allport, D. A., & Wylie, G. R. (1999). Task-switching: Positive and negative priming of task-set. In G. W. Humphreys, J. Duncan, & A. M. Treisman (Eds.), *Attention, space, and action: Studies in cognitive neuroscience* (pp. 273–296). Oxford, England: Oxford University Press.

Allport, D. A., & Wylie, G. R. (2000). Task-switching, stimulus-response bindings, and negative priming. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Cognitive control* (pp. 35–70). Cambridge, MA: MIT Press.

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Lebiere, C. (1998). *The atomic component of thought.* Mahwah, NJ: Erlbaum.

Baddeley, A. (1986). *Working memory.* Oxford, England: Oxford University Press.

Carlson, R. A., & Lundy, D. H. (1992). Consistency and restructuring in learning cognitive procedural sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 127–141.

De Jong, R. (2000). An intention–activation account of residual switch costs. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Cognitive control* (pp. 357–376). Cambridge, MA: MIT Press.

De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situation. *Acta Psychologica, 101,* 379–394.

Dreisbach, G., Haider, H., & Kluwe, R. H. (2001). *Preparatory processes in the task switching paradigm: Evidence from the use of probability cues.* Manuscript submitted for publication.

Gopher, D., Armony, L., & Greenshpan, Y. (2000). Switching tasks and attention policies. *Journal of Experimental Psychology: General, 129,* 308–339.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology* (Whole No. 89).

Lhermitte, F. (1983). "Utilization behaviour" and its relation to lesions of the frontal lobes. *Brain, 106,* 237–255.

Logan, G. D. (1985). Executive control of thought and action. *Acta Psychologica, 60,* 193–210.

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General, 129,* 4–26.

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1124–1140.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1423–1442.

Meiran, N. (2000). Reconfiguration of stimulus task-sets and response task-sets during task switching. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Cognitive control* (pp. 377–399). Cambridge, MA:MIT Press.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task-switching. *Cognitive Psychology, 41,* 211–253.

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory:

Role of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General, 106,* 226–254.

Norman, D. A., & Shallice, T. (1986). *Attention to action: Willed and automatic control of behavior.* In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1–18). New York: Plenum.

Owen, A. A., Roberts, A. C., Hodges, J. R., Summers, B. A., Polkey, C. E., & Robbins, T. W. (1993). Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or Parkinson's disease. *Brain, 116,* 1159–1175.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1–66.

Shallice, T. (1994). Multiple levels of control processes. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing.* (pp. 395–420). Cambridge, MA: MIT Press.

Sohn, M.-H., & Carlson, R. A. (1998). Procedural framework for simple arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1052–1067.

Sohn, M.-H., & Carlson, R. A. (2000). Effects of repetition and foreknowledge in task-set reconfiguration. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1445–1460.

Sohn, M.-H., Ursu, S., Anderson, J. R., Stenger, V. A., & Carter, C. S. (2000). The role of prefrontal cortex and posterior parietal cortex in task-switching. *Proceedings of National Academy of Science, 97,* 13448–13453.

Spector, A., & Biederman, I. (1976). Mental set and shift revisited. *American Journal of Psychology, 89,* 669–697.

Wylie, G. R., & Allport, D. A. (2000). Task switching and the measurement of "switch costs." *Psychological Research, 63,* 212–233.

Appendix

Accuracy Results

Experiment 1

A trial was counted as correct only when both Task 1 and Task 2 in the trial were correctly responded. In Experiment 1, the overall accuracy was .83. The main effect of practice was significant, $F(2, 62) = 22.04$, $p < .0001$, $MSE = 0.01$. Accuracy was lower in Session 1 (.79) than in Session 2 (.84) and Session 3 (.85), Newman-Keuls, $p < .01$. The main effect of RSI was also significant, $F(2, 62) = 5.04$, $p < .01$, $MSE = 0.02$. Accuracy was lower with 200-ms RSI (.81) than with 1,500 ms RSI (.84), Newman-Keuls, $p < .01$. Accuracy with 600-ms RSI (.83) was not statistically different from these. The main effect of foreknowledge was significant, $F(1, 31) = 16.81$, $p < .01$, $MSE = 0.01$. Accuracy was higher with foreknowledge (.84) than with no foreknowledge (.82). The main effect of transition was significant, $F(1, 31) = 72.88$, $p < .0001$, $MSE = 0.01$. Accuracy was higher with repetition trials (.86) than with switch trials (.80).

Several two-way interactions were significant. The main effect of RSI was different depending on session, $F(4, 124) = 2.88$, $p < .05$, $MSE = 0.01$. RSI effect was significant in Session 1, $F(2, 62) = 5.33$, $p < .01$, $MSE = 0.004$, but not significant in other sessions, $p > .05$. Foreknowledge effect interacted with RSI effect, $F(2, 62) = 5.14$, $p < .01$, $MSE = 0.003$. Foreknowledge effect was significant with 600-ms RSI (.03) and 1,500-ms RSI (.03), $t(31) = .0001$, and $t(31) = .01$, respectively, but not significant with 200-ms RSI (.004), $p > .6$. Transition effect also interacted with RSI, $F(2, 62) = 4.48$, $p < .05$, $MSE = 0.003$. Transition effect with 200-ms RSI (.07) was greater than with 1,500-ms RSI (.05), Newman-Keuls, $p < .05$. However, these means were not different from the accuracy with 600-ms RSI (.06). Also, the three-way interaction involving practice session, foreknowledge, and transition was significant, $F(2, 62) = 3.16$, $p < .05$, $MSE = 0.003$. However, simple effect analyses revealed no statistically significant differences of transition effect between foreknowledge and no foreknowledge in any sessions, $p > .09$. All of these interactions seem to imply that the difference between foreknowledge conditions or between task transitions became smaller when there was longer processing time (i.e., RSI) or when the task was more practiced (i.e., later session).

Experiment 2

In Experiment 2, overall accuracy was .80. Accuracy was higher in the second half (.82) than in the first half (.77), $F(1, 23) = 10.87$, $p < .01$, $MSE = 0.04$. Accuracy was higher with 3,000-ms (.80) and 5,000-ms (.80) RSIs than with 1,000-ms RSI (.78), $F(2, 46) = 3.62$, $p < .05$, $MSE = 0.01$. Accuracy was higher with foreknowledge (.81) than with no foreknowledge (.79), $F(1, 23) = 11.42$, $p < .01$, $MSE = 0.01$. Accuracy was also higher with task repetition (.82) than with task switch (.77), $F(1, 23) = 19.91$, $p < .01$, $MSE = 0.02$. The interaction between RSI and task transition was significant, $F(2, 46) = 4.57$, $p < .05$, $MSE = 0.01$. With task repetition, accuracy did not differ depending on RSI, $p > .9$. With task switch, accuracy was higher with 3,000-ms (.79) and 5,000-ms (.79) RSIs, than with 1,000-ms RSI (.74), $F(2, 46) = 6.77$, $p < .01$, $MSE = 0.002$, Newman-Keuls, $p < .01$. No other main effects or interactions were significant, $p > .10$.

Received July 20, 2000
Revision received March 14, 2001
Accepted May 25, 2001 ■