On the Limits of Advance Preparation for a Task Switch: Do People Prepare All the Task Some of the Time or Some of the Task All the Time?

Mei-Ching Lien
Oregon State University

Eric Ruthruff, Roger Remington, & James Johnston
NASA Ames Research Center

This study investigated the nature of advance preparation for a task switch, testing two key assumptions of De Jong's (2000) failure-to-engage theory were tested: (a) task-switch preparation is all-or-none, and (b) preparation failures stem from not utilizing available control capabilities. In all three experiments switch costs varied dramatically across the individual S-R pairs of the tasks - virtually absent for one S-R pair but large for other S-R pairs. These findings indicate that, across trials, task preparation was not all-or-none but rather was consistently partial (full preparation of some S-R pairs but not others). In other words, people do not prepare all of the task some of the time, but rather prepare some of the task all of the time. Experiments 2 and 3 produced substantial switch costs even though time deadlines provided strong incentives for optimal advance preparation. Thus, there was no evidence that people have a latent capability to fully prepare a task switch.

People are capable of performing an enormous range of tasks. To perform tasks relevant to current goals, it is necessary to impose “top-down” executive control over cognitive processing rather than simply react to the most salient stimulus. Although such executive control is clearly possible, it is subject to some limitations. One important limitation is that switching from one task to another incurs a cost in time and/or accuracy. Following the discovery of this switch cost by Jersild (1927), considerable research has been carried out to determine its implications for the nature of executive control (see Monsell & Driver, 2000).

One classical method used to study task switching is to compare performance in pure-task blocks (no switching required) to performance in alternating-task blocks (switching required). Response times (RTs) are typically longer in the alternating-task blocks than in the pure-task blocks, producing a switch cost (e.g., Allport, Styles, & Hsieh, 1994; Jersild, 1927; Spector & Biederman, 1976). A drawback of this method, noted by Rogers and Monsell (1995; see also Los, 1996), is that participants have to keep two stimulus-response (S-R) mappings in mind throughout the alternating-task blocks but not the pure-task blocks. Consequently, the cost of switching tasks is confounded with the cost of keeping two task sets active. To avoid this confound, Rogers and Monsell (1995) developed an alternating-runs paradigm (e.g., AABBAABB…), in which task-switch trials and task-repetition trials are intermixed within the same block. Another common method used to study task switching is to present a task cue prior to each trial, in an otherwise unpredictable task sequence, and examine how the switch cost is affected by the preparation time (e.g., Meiran, 1996).

Despite the variety of paradigms used to investigate task-switching, studies have generally yielded similar results. One of the most robust findings is that switch costs are larger when each stimulus is appropriate for performance of both tasks (“dual-affordance” or “bivalent” stimuli) than when each stimulus is appropriate for only one task (“single-affordance” or “univalent” stimuli). In a dual-affordance condition, each stimulus has the potential to automatically activate responses from both the currently relevant and currently irrelevant task sets. Thus, executive control is needed to prevent application of the irrelevant task set or, failing that, to resolve the conflict between multiple activated responses.

Dual-affordance switch costs are robust, remaining substantial even when (a) the task identity is provided in advance of each trial (e.g., Lien, Schweickert, & Proctor, 2003; Sohn & Carlson, 2000), and (b) the time interval between each response and the subsequent stimulus onset (response-stimulus interval; RSI) is relatively long (e.g., Meiran, 1996, 2000; Rogers & Monsell, 1995). The existence of a *residual* switch cost with ample opportunity to prepare in advance for a new task indicates that executive task-control ability is limited in some way.

Theoretical Interpretations of Residual Switch Costs

Several alternative, but not mutually exclusive, hypotheses have been proposed to explain why a residual switch cost occurs even with ample opportunity for preparation. Allport and his colleagues (Allport et al., 1994; Allport & Wylie, 2000) attributed costs on task-switch trials to task-set activation carried over from the previous trial (*task set inertia*) that cannot easily be suppressed by top-down preparation for a new task. Mayr and Keele (2000) hypothesized that the recently used task set is inhibited when switching to a new task and that this inhibition cannot easily be undone when the inhibited task later becomes relevant again (see also Hübner, Dreisbach, Haider, & Kluwe, 2003; Schuch & Koch, 2003). Rogers and Monsell (1995) argued that although some top-down task-set reconfiguration for a task switch can be carried out during the preparation interval, task-set reconfiguration cannot be completed without first processing a stimulus for the new task (see also Logan & Gordon, 2001; Meiran, 1996, 2000; Monsell, Yeung, & Azuma, 2000; Rubinstein, Meyer, & Evans, 2001).

Despite the differences among these competing accounts of switch costs, they share two important assumptions. First, they assume that some preparation for an upcoming task-switch occurs, but it is incomplete (compared to preparation for a task repetition). Second, they assume that residual switch costs reflect a fundamental (structural) inability to fully prepare for a task switch. As Gopher, Armony, and Greenshpan (2000) summarized, "whatever factors are involved, their influence is not amenable to voluntary, advanced preparation" (p. 311).

Failure-To-Engage Theory

In contrast to the prevailing structural theories of switch costs, De Jong (2000) proposed that people are, in fact, capable of preparing their cognitive systems to perform a task switch trial just as quickly and accurately as a task-repetition trial. To explain the existence of a switch cost, De Jong relied on a kind of performance/competence distinction. Although people have the competence to prepare fully for a task switch, their performance frequently does not reflect this competence. De Jong called this failure to utilize available control capabilities "goal neglect" (see also De Jong, Berendsen, & Cools, 1999). Among the factors that contribute to preparation failures, according to De Jong, are (a) weak goal-driven intention (e.g., lack of motivation), (b) weak environmental support (e.g., a lack of explicit task cues or a lack of clear feedback), and (c) special circumstances such as fatigue.

According to De Jong's failure-to-engage (FTE) theory, advance preparation is all or none. Thus, task-switching data obtained with a long RSI should contain a mixture of one subset of trials in which participants fully engage in advance task preparation and another subset of trials in which they completely fail to engage in advance task preparation. De Jong (2000) argued that the RT distribution for the subset of prepared switch trials should closely resemble that obtained on task-repetition trials (assumed to always be prepared). He also argued that the RT distribution for unprepared switch trials should resemble that obtained on short-RSI task-switch trials (assumed to always be unprepared). As a consequence, the RT distribution on long-RSI task-switch trials should be a mixture of the RT distribution for task-repetition trials and the RT distribution for short-RSI task-switch trials.

De Jong (2000) tested this prediction by computing cumulative distribution functions (CDFs) for task-repetition trials (collapsed across long and short RSIs), long-RSI task-switch trials, and short-RSI task-switch trials. These three CDFs are shown in Figure 1. Consistent with the FTE mixture hypothesis, the CDF for long-RSI switch trials approached the CDF for short-RSI task-switch trials at the high-percentile end of the distribution (long RTs) but approached the CDF for task-repetition trials at the low-percentile end of the distribution (fast RTs). Furthermore, De Jong found that a quantitative mixture model could provide a close fit to the observed CDFs when the probability of engaging in preparation was set to .51. Thus, although De Jong's model makes strong assumptions (e.g., that advance preparation is all-or-none), it provides an excellent fit to detailed aspects of his RT data.

The all-or-none preparation assumption of De Jong's (2000) FTE theory appears to be shared by a model proposed by Mayr and Kliegl (2000). While De Jong did not specify the nature of advance preparation, Mayr and Kliegl speculated that task preparation involves the retrieval of S-R mapping rules from long-term memory. They argued that preparation failures arise from failing to retrieve the S-R mapping rules from long-term memory. Mayr and Kliegl argued that "an all-or-none pattern of preparedness for an upcoming task easily follows from the retrieval failure account, possibly in combination with motivational constraints on preparation, as suggested by De Jong" (p. 1138).

The Present Study

De Jong's (2000) FTE theory differs from traditional task-switching theories in two key respects, both of which have important implications for theories of executive control. First, FTE theory proposes that advance preparation for a task switch is all-or-none across trials, whereas traditional task-switching theories assume some form of partial preparation. Second, FTE theory proposes that lack of advance preparation is due to failures to utilize available cognitive control capabilities (hereafter, *utilization failures*), whereas traditional task-switching theories assume a structural limitation. The present study will address both of these issues.

Is Advance Preparation All-or-None or Partial?

De Jong (2000) reported that the CDF for long-RSI task-switch trials was fit well by a mixture of completely prepared states and completely unprepared states. This finding appears to support his all-or-none preparation assumption and therefore constitutes a strong challenge to competing theories in which advance preparation for a task switch is typically partial. In particular, this result appears to undermine Rogers and Monsel's (1995) claim that participants complete only part of task preparation in advance of a stimulus, leaving the rest to be completed after stimulus onset. The results also appear to undermine Rubinstein et al.'s (2001) proposal that people "shift goals" in advance of the stimulus, but do not load the specific S-R mapping rules until after the stimulus has been identified (see also Goschke, 2000). Both of these models clearly predict that advance preparation for each trial should be partial, rather than all-or-none.

Although De Jong's (2000) analysis might appear to rule out all partial preparation models of task switching, at least one important sub-case is not ruled out. In what follows, we describe a plausible hypothesis in which the preparatory state is partial on every trial, but the outcome of that preparatory state is all-or-none depending on which stimulus is presented. In addition to accounting for De Jong's findings, this *partial mapping preparation hypothesis* describes what people do to prepare for a task switch more concretely than most previous theories.

The partial mapping preparation hypothesis assumes that people always engage in preparation for an upcoming task switch, at least under favorable conditions (e.g., with ample preparation time). This deliberate preparation, however, is limited because participants can prepare for only a subset of S-R pairs on a typical trial. In De Jong's red/blue task, for example, participants might have been prepared to respond to red stimuli (with the left key), but not to blue stimuli (or vice versa). Participants would respond quickly to the prepared subset of stimuli (e.g., red letters), but slowly to the unprepared stimuli (e.g., blue letters), producing a mixture of results. Thus, this hypothesis provides a straightforward explanation for why De Jong's data fit an all-or-none mixture model. In this case, the mixture results only from the "luck of the draw" – whether the presented stimulus happened to be one that was prepared or one that was unprepared. Although the data reflect a mixture of outcomes, the underlying preparatory state is actually relatively constant across trials.

Although the proposed partial mapping preparation hypothesis has the potential to explain De Jong's (2000) CDF results, it has not yet been directly tested. One might hope to test this hypothesis by examining whether certain S-R pairs show little switch cost while other S-R pairs show unusually large switch costs. However, we are aware of no previous studies that reported switch costs as a function of individual S-R pairs. Furthermore, in most previous task-switching studies, each task mapped only two stimulus categories (e.g., even and odd) to two responses (e.g., left and right keypresses; De Jong, 2000; Goschke, 2000; Hübner, Futterer, & Steinhauser, 2001; Lien & Ruthruff, 2004; Nieuwenhuis & Monsell, 2002; Rogers & Monsell, 1995; Sohn & Carlson, 2000). It is not clear in such cases that one particular S-R pair would usually be prepared; rather, the preparation strategy is likely to vary across and within participants, preventing a clean test of the partial mapping preparation hypothesis. A further problem is that many previous studies mixed different RSI values (short and long) during the experiment. Experiments with short RSIs, where there is insufficient time even for partial preparation, might discourage participants from preparing as much as possible on long-RSI trials (see Rogers & Monsell, 1995).

Because no previous data allow a test between De Jong's (2000) all-or-none preparation hypothesis and our partial-mapping hypothesis, the present experiments were designed to provide such a test. We used tasks that mapped three stimulus categories (e.g., the colors red, green, and blue) onto three response categories (left, middle, and right response keys)[1]. Instructions described the S-R pairs in the order of response key position: left, middle, right. For ease of discussion, we refer to these S-R pairs as *the first S-R pair*, *the second S-R pair*, and *the third S-R pair* (see Figure 2). We hypothesized that participants would think of the S-R pairs in the normal left-to-right order of English reading, reinforced by using the same order of presentation in the instructions. Thus, we expected participants to prepare for the S-R pairs in order from left (the first S-R pair) to right (the third S-R pair).

Based on the partial mapping preparation hypothesis, and the assumption of a predominantly left-to-right preparation order, we predicted that advance preparation should routinely be devoted to the first S-R pair, resulting in little or no switch cost for this pair. The third S-R pair, meanwhile, should rarely be prepared, producing a large switch cost. The switch cost for the second S-R pair should fall between these two extremes, depending on how often participants are capable of preparing for two S-R pairs rather than just one. According to FTE theory, in contrast, participants should be completely unprepared for all S-R pairs on some proportion of task-switch trials, resulting in switch costs for all three S-R pairs.

Is The Lack of Advance Preparation Due to Utilization Failure or a Structural Limitation?

The second issue addressed in the present study is the nature of human limitations in advance task preparation. Traditional task-switching theories imply that the lack of advance preparation is due to a fundamental inability to complete preparation (a structural limitation). De Jong (2000), however, argued that people are in fact fully capable of completing advance preparation for a task switch, but fail to utilize this capability on some proportion of trials. This utilization failure need not arise from a conscious decision; it could arise simply from not exerting enough positive effort or even just forgetting to prepare at all.

To address this issue, Nieuwenhuis and Monsell (2002) adopted Rogers and Monsell's (1995) experimental design but made three major modifications to provide incentives for advance preparation. First, they provided detailed performance feedback after each block and encouraged participants to strive for continuous performance improvements. Second, they used a reward system, in which participants were given a monetary bonus for fast and accurate performance. Third, they reduced the number of trials in each block (but kept the total number of trials the same as in Rogers and Monsell) to minimize fatigue-induced failures to engage in advance preparation. Although these efforts reduced both overall RT and the switch cost relative to that found by Rogers and Monsell, a substantial switch cost remained (69 ms) and the mixture model still provided a reasonably good fit. Nieuwenhuis and Monsell concluded that switch costs are caused by more than just a motivational limitation on participants' ability to fully engage in advance preparation.

Although Nieuwenhuis and Monsell's (2002) study constitutes a useful start in assessing whether switch costs result from utilization failure or a structural limitation, it is far from decisive. The main reason is that Nieuwenhuis and

Monsell followed the long tradition of carrying out task-switching studies with the standard RT procedure. In this procedure, instructions usually emphasize both speed and accuracy (e.g., "please respond quickly and accurately") and feedback on these performance measures are provided at the end of each block. In the case of Nieuwenhuis and Monsell, mean RTs and error rates were graphically presented at the end of each block of trials, but there was no RT feedback after each trial. Without immediate RT feedback on each trial, participants might not realize that advance preparation is critical for performance and is important enough to justify expending additional effort (see De Jong, 2000). A related point is that the standard RT procedure does not penalize a prolonged RT on a given trial; in fact, such delays typically are not acknowledged in any way. Although participants are sometimes required to respond within a fixed "time-out" period on each trial, this time period is generally so long (e.g., 5 seconds in Nieuwenhuis & Monsell) that feedback almost never occurs. This analysis suggests that Nieuwenhuis and Monsell's procedure, like most traditional task-switching experiments, provided far from optimal incentives to encourage participants to apply their full cognitive control capabilities.

To shed light on whether switch costs reflect structural limitations or utilization failure, we developed a new procedure for assessing switch costs. The basic strategy was to set a tight time deadline so that a switch cost would cause participants to miss that deadline, immediately triggering failure feedback (see Ruthruff, Johnston, & Remington, 2003, for an application of a similar logic to the study of dual-task interference). Unlike the standard RT procedure, this deadline procedure provides clear and immediate negative consequences (an "error" or "too slow" message) for slow responses due to poor preparation. The desire to avoid such negative feedback should provide strong incentives for participants to prepare for a task switch to the full extent of their capabilities.

The present Experiment 1 used the traditional RT procedure, whereas Experiments 2 and 3 used the novel time-deadline procedure. Obtaining robust switch costs even with the time-deadline procedure would considerably strengthen the case for a fundamental inability to completely prepare for a task switch (as assumed by our partial mapping preparation hypothesis). Obtaining little or no switch cost with the time-deadline procedure would support the FTE claim that people are capable of completing advance preparation for a task switch.

Experiment 1

The purpose of Experiment 1 was to determine, using the standard RT procedure, whether advance preparation is all-or-none (as suggested by the FTE theory) or partial (as suggested by the partial mapping preparation hypothesis). The design used three S-R pairs for each task, and the traditional RT procedure, with instructions to respond quickly and accurately (i.e., no time deadline). As in De Jong (2000), we used dual-affordance tasks that clearly require a high degree of executive control. The stimulus on each trial was a colored shape; participants responded either to its color (the color task) or to its shape (the shape task).

Following the alternating-runs paradigm of Rogers and Monsell (1995), an AABB task sequence was used and the stimulus location provided a straight-forward task cue. Each stimulus appeared in one of four boxes arranged in a two-by-two grid, as shown in Figure 3. The first stimulus appeared in the upper left box and each subsequent stimulus appeared in the box located immediately clockwise from the previous one. One task was to be performed on all stimuli displayed in the top two boxes and the other task was to be performed on all stimuli displayed in the bottom two boxes (the assignment was counterbalanced across participants).

Several steps were taken to encourage advance preparation on task-switch trials. First, the instructions explicitly encouraged participants to prepare for each trial in advance so that they could respond as quickly and accurately as possible. Second, because short RSIs discourage advance preparation, we used a long RSI on every trial. Third, we set the RSI at 2 seconds (well beyond the point at which switch costs stop decreasing). Fourth, we always selected an irrelevant attribute associated with a response that was incompatible with the correct response, ensuring that failures to switch task set would always result in an erroneous response with error feedback.

According to our partial mapping preparation hypothesis, switch costs should depend strongly on the S-R pair. Specifically, there should be little or no switch cost for the first S-R pair, but a large switch cost for the third S-R pair (and an intermediate result for the second S-R pair). However, if residual switch costs are due to occasional utilization failure (all-or-none preparation), then there is no obvious reason for switch costs to depend strongly on the S-R pair.

Method

Participants. A total of 18 participants from colleges and universities surrounding the National Aeronautics and Space Administration (NASA) Ames Research Center participated in exchange for extra course credit. All participants had normal or corrected-to-normal vision.

Apparatus and stimuli. The stimuli were presented on IBM-compatible microcomputers connected to SONY Trinitron monitors, housed in a dedicated, sound-attenuating booth. A 12 cm × 12 cm frame of four boxes was presented in the center of the screen (see Figure 3). On each trial, a colored object (a triangle, square, or diamond shape filled with the color red, green, or blue) appeared in one of these boxes. Each side of the diamond and square was 1.5 cm long. The horizontal side (bottom) of the triangle was 1.5 cm and the other two sides were 2.0 cm.

Design and procedure. Participants performed either the color or the shape task on each trial. For the color task, participants determined whether the object was red, green, or blue and responded by pressing the "M", "<", or ">" key, respectively. For the shape task, participants determined whether the object was triangle, diamond, or square and responded by pressing the "M", "<", or ">" key, respectively. Participants pressed these three keys (M, <, and >) with the index, middle, and ring fingers of their right hand, respectively (see Figure 2). The instructions described all S-R mappings for one task in the order of the leftmost response, the middle response, and then the rightmost response and all S-R mappings for the other task in the same order (to induce participants to prepare for the S-R pairs in the same order). The order of the mapping instructions for each task was counterbalanced across participants. On each trial, the relevant and irrelevant attributes were chosen at random, with the restriction that they always corresponded to different responses.

Following the procedure of Rogers and Monsell (1995), the first stimulus of each block appeared in the top-left box. Each subsequent stimulus appeared in the box located immediately clockwise from the previous one. Depending on the location of the stimulus, participants were to perform either the color discrimination task or the shape discrimination task. For half of the participants, the top two locations were assigned to the colr task and the bottom two locations were assigned to the shape task (thus, the task sequence was color-color-shape-shape-color-color…). For the other half of participants, the assignment was reversed. The task cues "COLOR" and "SHAPE" appeared next to the particular locations associated with that task for the first 3 blocks of practice.

To motivate participants, the tasks were described within the context of a game. Participants were told that they were controlling a starship (represented by the center circle in Figure 3) being attacked by missiles. Participants were instructed to fire their laser and destroy the enemy missiles quickly and accurately. On each trial, the starship's laser pointed to the location of the upcoming missile. In addition, a plus sign appeared for 1,000 ms in the center of that box. A colored shape then appeared 500 ms after the offset of the plus sign. It remained on the screen until the participant made a response or until 4 seconds had elapsed. If the response was correct, the stimulus was replaced by a yellow "smiley face" symbol for 300 ms. If the response was incorrect, the starship exploded (replaced by a pattern of white dots that expanded outward in successive frames) for 250 ms and then was replaced by a frowning face for 600 ms; the frowning face was accompanied by an error beep. The next trial began 500 ms following the offset of this positive or negative feedback. Consequently, the total time interval from the offset of the feedback in the previous trial to the onset of the stimulus in the next trial was 2 seconds.

Participants performed 12 blocks of 40 trials, the first four of which were considered practice. The experiment lasted approximately 25 minutes. Participants were told that both the speed and accuracy of responding were very important (the standard instructions used in task-switching studies). They were also encouraged to take a brief break before beginning the next block. At the end of each block, participants received a summary of their average RT and average accuracy for that block.

Data Analyses. The first four trials in each experimental block (one cycle through the 2×2 stimulus grid), which served as warm-up trials, were omitted from the analyses. For RT analyses, trials were also omitted if the current or previous response was an error. RTs outside the range of 200 ms to 2,300 ms were treated as outliers, which led to the elimination of an additional .3% of the trials.

Single-task studies have consistently found that choice responses are much faster when the relevant stimulus repeats (and thus the correct response also repeats) than when it changes (e.g., Bertelson, 1961, 1965; Campbell & Proctor, 1993; Pashler & Baylis, 1991a, 1991b). This stimulus/response repetition effect may be caused by "creation and transient strengthening of links that shortcut the response-selection stage, by translating directly from fairly early stimulus representations all the way to fairly specific responses" (Pashler & Baylis, 1991b, p. 46). For instance, participants might simply repeat the last response, without retrieving from memory the response mapped to the relevant stimulus. Because this short-cut is available on task-repetition trials but not on task-switch trials, it could by itself produce a task switch cost. Therefore, to isolate task repetition effects from stimulus repetition effects, previous studies have generally either prevented relevant stimulus repetitions from occurring in the first place (e.g., Rogers & Monsell, 1995, p. 212), or allowed them to occur but eliminated them from analysis after the fact (e.g., Lien & Ruthruff, 2004; Ruthruff, Remington, & Johnston, 2001). We chose the latter approach in the present study to preserve the natural probabilities for the occurrence of stimuli and responses (i.e., one in three). Relevant stimulus repetitions are, by definition, not possible on task-switch trials and therefore do not need to be excluded. As discussed below, other forms of stimulus and response repetition are possible (e.g., the irrelevant stimulus on one trial can become relevant on the next trial), but none permit short-cutting response selection and none produce substantial performance benefits.

Task type (color versus shape) had little effect on performance and did not consistently interact with other factors. Consequently, this variable was not included as a factor in the final data analyses reported here. Data were analyzed as a function of task transition (repetition versus switch) and S-R pair (the first, second, and third). Significant interactions between task transition and S-R pair were followed up with separate analyses to clarify the relations. An alpha level of .05 was used to determine statistical significance.

Results

RT Data. Table 1 shows the RT and proportion of error (PE) data for the task-repetition and task-switch trials. For RT data, there was a significant switch cost of 68 ms, $F(1, 17) = 22.94$, $p < .001$, $MSE = 5{,}426$. There was also a main effect of S-R pair, $F(2, 34) = 14.59$, $p < .001$, $MSE = 8{,}150$; mean RT was shortest for the first S-R pair and longest for the third S-R pair.

Switch costs varied significantly across S-R pairs, $F(2, 34) = 11.91$, $p < .001$, $MSE = 4{,}346$. To follow up on this significant interaction, five separate ANOVAs were used to test for simple main effects of S-R pair for each task transition (task repetition and task switch) and simple main effects of task transition for each S-R pair (the first, second, and third). The effect of S-R pair was significant for task-switch trials, $F(2, 17) = 20.60$, $p < .001$, $MSE = 7{,}696$; mean RT was shortest for the first S-R pair, intermediate for the second S-R pair, and longest for the third S-R pair. In contrast, the effect of S-R pair was not significant for task-repetition trials, $F(2, 17) = 2.98$, $p = .10$, $MSE = 4{,}800$. The switch cost was not significant for the first S-R pair, $F(1, 17) < 1$, but was significant for the second and third S-R pairs, $Fs(1, 17) \geq 23.61$, $ps < .001$, $MSEs \leq 4{,}949$. The switch cost was -20 ms for the first S-R pair but was 107 ms and 117 ms for the second and third S-R pairs (see Table 1).

Error Data. There was also a significant switch cost on PE, $F(1, 17) = 6.73$, $p < .05$, $MSE = .0014$; the switch cost on PE was .02. Although S-R pair had no main effect on PE, it did modulate the switch cost on PE, $F(2, 34) = 3.44$, $p < .05$, $MSE = 0.0010$. In the follow-up analyses, the effect of S-R pair was significant for the task-repetition trials, $F(2, 34) = 3.40$, $p < .05$, $MSE = 0.0013$, but not for the task-switch trials, $F(2, 34) = 2.16$, $p = .13$, $MSE = 0.0015$ (see Table 1). The switch cost was significant for the second S-R pair, $F(1, 17) = 20.04$, $p < .001$, $MSE = 0.0007$, but not for the first and third S-R pairs, $Fs(1, 17) \leq 1$; the switch costs were .01, .04, and .01 for the first, second, and third S-R pairs.

Stimulus/Response Repetition. The preceding analyses excluded task-repetition trials in which the relevant stimulus repeated because they permit participants to short-cut response selection. A separate set of data analyses were carried out on task-repetition trials to assess stimulus repetition effects. Each analysis compared the condition with no form of stimulus/response repetition (the "no repetition" condition) to one of the conditions with some form of repetition. Table 2 shows mean RT for each of these conditions.

Repeating the irrelevant stimulus, without repeating the relevant stimulus, reduced RT by a non-significant 11 ms, $F(1, 17) < 1$. In contrast, repeating the relevant stimulus, without repeating the irrelevant stimulus, reduced mean RT by 62 ms, $F(1, 17) = 4.71$, $p < .05$, $MSE = 103{,}644$. We argue that when the relevant stimulus repeats, participants might simply repeat the response from the previous trial (short-cutting the usual response selection process). The reduction in RT was even larger (128 ms) when both the relevant and irrelevant stimuli repeated, $F(1, 17) = 23.48$, $p < .001$, $MSE = 18{,}976$. Repetition of the entire stimulus ensemble might enhance recognition that the relevant stimulus repeated and therefore encourage use of the short-cut strategy. The effect of repeating either the relevant stimulus or the entire stimulus ensemble did not interact with S-R pair, $Fs(2, 34) < 1$. Thus excluding the stimulus repetition trials from the main data analyses reported above was not responsible for the main finding that switch costs were much smaller for the first S-R pair than for the second and third S-R pairs.

On task-switch trials, the relevant stimulus (e.g., red for the color task) cannot repeat, so there are no trials that allow the participants to short-cut response selection by repeating the last response. Nevertheless, three other types of stimulus or response repetitions are possible – (1) the correct response can repeat, (2) the irrelevant stimulus can be identical to the relevant stimulus from the previous trial, and (3) the relevant stimulus can be identical to the irrelevant stimulus from the previous trial (potentially causing negative priming; see Fox, 1995). Our design allowed response repetitions to occur only in the absence of stimulus repetition[2], but did allow both types of stimulus repetition to occur on the same trial (e.g., when performing the color task on a red triangle, then performing the shape task on a red triangle). See Table 2 for a summary of the possible types of repetition and the corresponding mean RTs.

Because the main data analyses excluded relevant stimulus/response repetitions on task-repetition trials, one could argue that they also have excluded any form of stimulus or response repetitions on task-switch trials. To test whether these types of repetition strongly influence task switch performance, we conducted additional data analyses on task-switch trials only. Each analysis compared the "no repetition" condition to one of the conditions with some form of repetition (see Table 2). The effect of repeating the response (cost of 7 ms) was not statistically significant, $F(1, 17) < 1$. There was a modest but significant benefit (44 ms) when the current irrelevant stimulus matched the previous relevant stimulus, $F(1, 17) = 8.87$, $p < .01$, $MSE = 53{,}668$. This effect failed to replicate in Experiments 2 and 3, so we will not attempt to interpret it. There was a small, non-significant "negative priming" cost (29 ms) when the current relevant stimulus was the same as the previous irrelevant stimulus, $F(1, 17) = 1.49$, $p = .2395$, $MSE = 14{,}649$. Repeating the entire stimulus ensemble on task-switch trials (where the current irrelevant stimulus was the previous relevant stimulus and the current relevant stimulus was the previous irrelevant stimulus) produced no benefit relative to the no-repetition condition; in fact, it produced a non-significant cost of 5 ms, $F(1, 17) < 1$. As in the data analyses for the task-repetition trials, none of the repetition conditions interacted with the S-R pair, $Fs(2, 34) < 1$. In summary, the presence of stimulus and/or response repetition on task-switch trials had little main effect and did not interact with the S-R pair. Consequently, our main findings do not depend on decisions about what types of stimulus or response repetition to include or exclude.

Discussion

Even though we used a constant, long RSI (2 secs) and a predictable task sequence, we observed substantial switch costs on both RT (68 ms) and PE (.02). These findings suggest that even when participants have ample time to prepare for a task switch, their preparation is imperfect. One novel finding of this experiment is that switch costs depend very strongly on the S-R pair. The switch cost on RT was – 20 ms for the first S-R pair[3], 107 ms for the second S-R par, and 117 ms for the third S-R pairs. The follow-up analyses of simple main effects indicated that this pattern occurred because RT increased strongly from the first to the third S-R pair on task-switch trials but not on task-repetition trials (see Table 1). This pattern of results supports the hypothesis that task repetition (bottom-up) prepares the entire task set, whereas executive mechanisms (top-down) prepare only one or two individual S-R pairs. In other words, top-down task preparation on individual trials is typically incomplete, as suggested by our partial mapping preparation hypothesis.

These results are inconsistent with the all-or-none preparation assumption of De Jong's (2000) FTE theory. According to this theory, task-switch trials consist of a mixture of trials where participants completely engage in advance preparation (for all S-R pairs) and trials where they completely fail to engage in preparation. The latter subset of trials (failures to engage) should have produced switch costs regardless of what stimulus category happened to be presented on those trials. In other words, substantial switch costs should have been observed for all three S-R pairs. Contrary to this prediction, we observed no switch cost for the first S-R pair. To explain our findings, the FTE theory would need to add the additional assumption that the first S-R pair is somehow immune to switch costs, even when completely unprepared. Contrary to this assumption, however, our pilot studies showed large switch costs for the first S-R pair at short RSIs, where participants presumably do not have sufficient time to complete preparation for any of the S-R pairs.

Cumulative Distribution Functions (CDFs)

By examining the CDFs for task switch and task-repetition trials, it is possible to roughly estimate (for each S-R pair) the proportion of prepared and unprepared trials. If task-switch trials are always prepared, then the task repetition and task switch CDFs should be very similar at all percentiles. If task-switch trials are never fully prepared, then the task switch CDF should be shifted to the right of the task repetition CDF at all percentiles. If task-switch trials contain a mixture of prepared and unprepared trials, however, then the CDFs should be similar at small percentiles but should diverge at large percentiles (as noted by De Jong, 2000).

To compute CDFs, the RTs for each condition of each participant were rank-ordered and partitioned into deciles (i.e., into 10 bins). The RTs for each decile were then averaged across participants. Figure 4 (top panel) shows that for the first S-R pair, the CDFs for the task-switch trials and the task-repetition trials were nearly identical. This result suggests that, on task-switch trials, participants were always fully prepared for the first S-R pair. For the third S-R pair, however, the CDF for the task-switch trials was shifted to the right of the CDF for the task-repetition trials at all percentiles (see Figure 4, bottom panel). This result suggests that, on task-switch trials, participants were rarely (if ever) prepared fully for the third S-R pair, despite the long RSI. For the second S-R pair, the CDF for the task-switch trials overlapped that for the task-repetition trials at small percentiles, but diverged at large percentiles (see Figure 4, middle panel). This result suggests that, on task-switch trials, participants were sometimes prepared for the second S-R pair and sometimes unprepared[4].

Interestingly, the evidence for a mixture of prepared and unprepared states for the second S-R pair could easily be explained by the FTE theory. However, this theory predicts that the same mixture should occur for the first and third S-R pairs as well, contrary to our observations. Whereas FTE theory cannot explain the pattern of CDFs, our partial mapping preparation hypothesis provides a straightforward explanation. According to our hypothesis, participants typically prepared fully for the first S-R pair[5], sometimes prepared for the second S-R pair, and rarely prepared fully for the third S-R pair.

Having proposed that there is a mixture of prepared and unprepared S-R pairs on task switches, it is worth considering whether the same might also be true on task-repetition trials. In the case of task repetitions, the prepared S-R pair might be the one participants just performed, rather than the first S-R pair (thus explaining why these trials are so fast)[6]. The attraction of this simple hypothesis is that it seeks to explain the data without invoking different mechanisms for task switch and task-repetition trials. Unfortunately, the facts argue against it. First, the benefit of stimulus repetition has been found even in cases where the task was not expected to repeat (Ruthruff et al., 2001), suggesting that it does not reflect deliberate preparation. More importantly, this hypothesis predicts that performance for unprepared task switches (the second and third S-R pairs) should be similar to performance for unprepared task repetitions (the not recently used S-R pairs). On the contrary, these supposedly unprepared task-repetition trials were much faster than the supposedly unprepared task-switch trials (see Tables 1-4). The data support the hypothesis that performing a task leads to a heightened state of preparation not just for the recently used S-R pair but for all the other S-R pairs of that task as well.

Experiment 2

Experiment 2 was designed to determine whether the relatively poor preparation on task-switch trials is due to utilization failures or a structural limitation. As noted earlier, the RT procedure used in previous task-switching studies (as well as our Experiment 1) does not provide strong incentives for participants to prepare in advance on task-switch trials. In this procedure, a failure to prepare for a task switch would result in only a modest delay in RT. There is no feedback regarding this delay and there is no strong incentive for participants to prevent the delay. Consequently, it is plausible that the apparent lack of preparation (i.e., switch costs) was not due to a fundamental inability but rather to the presence of weak incentives.

To provide much stronger incentives for optimal performance, we developed a new deadline methodology. Each response must be made before a predetermined time deadline has expired; failures to do so trigger immediate negative feedback, even if the response was correct. The deadline time in Experiment 2 was adjusted using a tracking method aimed at an 80% success rate on task-repetition trials (so that there was just enough time to complete the task on most trials, but not much more). The same deadline time was then applied to task-switch trials as well, allowing no extra time for a switch cost. Consequently, if participants did not fully prepare in advance for a task switch, they would tend to fail, either by responding too late or by making an error. Assuming that participants do not like to fail, this method provides strong incentives for participants to prepare as much as possible. Unlike the standard RT paradigm, this paradigm allows us to more confidently attribute any observed switch costs to a structural preparation limitation rather than utilization failure.

Although time-deadline procedures have not previously been used to study task switching, they have been used to address other issues (e.g., speed-accuracy tradeoffs). These studies generally signaled the end of the deadline with an auxiliary stimulus such as a tone (e.g., Wickelgren, 1977). However, processing of the deadline signal places an extra burden on the participant. Another problem is that, as participants process the stimulus, there is no external indicator of how much time remains before the deadline (once the deadline signal occurs, it is too late to do anything about it). Therefore, participants presumably must set their own internal deadlines, which might be difficult and error prone (see Ruthruff, 1996). Furthermore, the relative desirability for responding in time at low accuracy versus responding late at high accuracy is unclear.

One way to remedy these problems is to create tasks with an inherent time deadline (e.g., shooting a target before it is out of sight). In the present implementation, we presented colored shapes in the outer corner of each box, which then moved steadily towards a circular icon in the center of the screen. Failures to respond to the colored-shape before it hit the center icon triggered immediate error feedback.

To further increase participants' motivation, and make this experiment more game-like, we modified the scenario described in Experiment 1. Participants were told that "You are the only crew member left on the starship *Enterprise* [represented by the center circle in Figure 3] and are surrounded by four enemy ships that are firing missiles to destroy your ship. In order to stop the enemy and save the ship, you have to battle against time. You must fire your laser before the enemy's weapon hits your ship. If you succeed (press the correct key for that weapon type BEFORE it hits your ship), you will see a smiley face on the screen. However, you get no credit for a late response because your ship has already been damaged. Likewise, you get no credit for making an early response if it is the wrong response – your ship will still be damaged because you do not have a second chance." This scenario makes it clear that late responses are pointless. Participants must respond correctly to the stimulus before it destroys the starship icon or fail.

Instead of measuring RT and PE, as in the standard RT paradigm, we measured success rate – the proportion of responses that were correct and in time (before the deadline). We then measured switch costs by subtracting success rates in the task-switch trials from success rates in the task-repetition trials. If the sole cause of switch costs in the RT procedure of Experiment 1 was utilization failure, then providing strong incentives for advance preparation might eliminate switch costs. However, if the cause of switch costs was a structural limitation (as assumed by the partial mapping preparation hypothesis), then switch costs should still be evident and should still depend strongly on the S-R pair.

Method

Participants. There were 24 participants, drawn from the same participant pool as in Experiment 1. None had participated in the previous experiment.

Apparatus, stimuli, and procedure. The tasks, stimuli, and equipment were the same as in Experiment 1 except that a deadline procedure was used to encourage participants to prepare as much as possible. Following the offset of the fixation sign, the colored shape appeared in the outer corner of the box and began moving towards the center of the screen (towards the starship). Participants were told to respond to this stimulus (the "missile") by pressing the appropriate response key before the missile reaches the center and damages the starship. Late responses were never rewarded: even if correct, late responses still resulted in damage to the starship and the associated negative feedback. At the end of each block, participants received a summary of their success rate for destroying the enemy's missiles during that block. They were also encouraged to keep improving their success rate in order to save their starship.

The initial deadline was 3 seconds, so that participants could focus on learning the S-R mappings. The deadline then declined steadily over the first two practice blocks until it reached a terminal value of 1 second. The deadline was adjusted thereafter using a tracking procedure, separately for each task, to achieve an average success rate of approximately .8 for task-repetition trials. The deadline for a task was increased whenever participants failed (responded late or inaccurately) on a task-repetition trial for that task. The deadline for a task was decreased whenever participants succeeded on the previous two task-repetition trials for that task (although the deadline for a task was not allowed to increase twice in a row). The size of the increase/decrease declined across blocks, so that smaller and smaller adjustments were made later in the experiment (on the assumption that the actual deadline was already close to the proper deadline)[7].

Results

As in Experiment 1, we excluded from analysis the first four trails (warm-up trials) in each experimental block. We also excluded any task-repetition trials where the stimulus was the same as the previous relevant stimulus. Overall, participants responded before the deadline on 92% of trials, suggesting that they understood the requirements of the deadline paradigm and adapted quickly (rarely responding too late). Responses made before the deadline were correct 76% of the time. This result suggests that, in order to respond before the deadline, participants were often forced to make a guess based on incomplete response information. The mean deadline was 654 ms (652 ms for the color task and 657 ms for the shape task), ranging from 536 ms to 947 ms across participants.

Becaue the same deadline was applied to both task-repetition and task-switch trials, any lack of advance preparation on task-switch trials should be reflected in a decrease in success rates for task-switch trials relative to task-repetition trials. Our primary interest in the time-deadline procedure, therefore, was the switch cost in the rate of success (correct response in time). Of secondary interest was the switch cost in RT. Note that in the time-deadline procedure RT depends not nly on the time to completely perform a task but also the time deadline. Consequently, caution is required when interpreting switch costs on RT. Switch costs in RT are useful, however, for establishing whether any switch costs on success rate could be due to a speed-accuracy tradeoff.

Success Rates. Table 3 shows the success rates (in proportion) for the task-repetition and task-switch trials. Success rate was higher for task-repetition trials than for task-switch trials, producing an overall switch cost of .084, $F(1, 23) = 37.40$, $p < .001$, $MSE = 0.0068$. There was a main effect of S-R pair, $F(2, 46) = 22.39$, $p < .001$, $MSE = 0.0131$; success rates decreased from the first S-R pair to the third S-R pair.

Switch costs varied significantly across S-R pairs, $F(2, 46) = 5.13$, $p < .01$, $MSE = 0.0092$. In the follow-up analyses of simple main effects, the effect of S-R pair was significant for both task-repetition trials, $F(2, 46) = 5.09$, $p < .05$, $MSE = 0.0115$, and task-switch trials, $F(2, 46) = 26.21$, $p < .001$, $MSE = 0.0108$. Success rate decreased from the first S-R pair to the third S-R pair for both task-repetition and task-switch trials, with the decrease being much more pronounced for the task-switch trials (see Table 3). The switch cost was not significant for the first S-R pair, $F(1, 23) < 1.0$, but it was significant for both the second S-R pair, $F(1, 23) = 8.57$, $p < .01$, $MSE = 0.0076$, and the third S-R pair, $F(1, 23) = 40.82$, $p < .001$, $MSE = 0.0067$. The switch costs on success rate were .027, .076, and .148 for the first, second, and third S-R pairs, respectively.

RT for Success Trials. Success trials, where responses were correct and in time, were also submitted for further data analyses on RT. As in the success rate analyses, there was a significant switch cost, $F(1, 23) = 6.15$, $p < .05$, $MSE = 376$; RT was 493 ms for the task-repetition trials and was 501 ms for the task-switch trials. Thus, the switch cost in RT was 8 ms. There was also a main effect of S-R pair, $F(2, 46) = 11.25$, $p < .001$, $MSE = 9{,}081$; RTs were 482, 499, and 509 for the first, second, and third S-R pairs, respectively. Switch costs varied significantly across S-R pairs, $F(2, 46) = 11.88$, $p < .01$, $MSE = 345$. For the task-repetition trials, RT was similar for all S-R pairs (488, 492, and 498 ms for the first, second, and third S-R pairs, respectively). For the task-switch trials, RT increased from 476 ms for the first S-R pair, to 506 ms for the second S-R pair, and to 521 ms for the third S-R pair. In other words, the switch costs on RT were -12, 14 ms, and 23 ms for the first, second, and third S-R pairs, respectively. This pattern was similar to (and in the same direction as) the switch cost on success rate across S-R pairs, indicating that the switch cost on success rate was not due to a speed-accuracy tradeoff.

Stimulus/Response Repetition. As in Experiment 1, we conducted separate data analyses to evaluate the effect of different types of repetitions on success rates for both task-repetition and task-switch trials. Table 4 shows the mean success rates for each type of repetition condition. The pattern of results is similar to that of Experiment 1. Repeating the irrelevant stimulus, without repeating the relevant stimulus, had little effect on success rates relative to the no-repetition condition, $F(1, 23) < 1$. However, repeating the relevant stimulus did substantially increase success rates both when the irrelevant stimulus was not repeated (an increase of .083), $F(1, 23) = 10.12$, $p < .01$, $MSE = 0.0249$, and when the irrelevant stimulus was also repeated (an increase of .142), $F(1, 23) = 59.28$, $p < .001$, $MSE = 0.0124$. These results are in consistent with the hypothesis that relevant stimulus repetition encourages participants to use the response-selection short-cut strategy. The effects of repeating either the relevant stimulus itself or the entire stimulus ensemble (i.e., both relevant and irrelevant stimuli) did not interact with S-R pair, $F\text{s}(2, 46) \leq 2.32$, $p\text{s} > .05$, replicating Experiment 1. Thus, the pattern of results (small switch costs for the first S-R pair, large switch costs for the second and third S-R pairs) would have been obtained even if we had not excluded trials with repetition of the relevant stimulus.

As in Experiment 1, we conducted the additional data analyses to evaluate whether task switch performance was influenced by stimulus or response repetition. Compared to the no-repetition condition, repeating the response had no significant effect on success rate, $F(1, 23) = 3.91$, $p > .05$, $MSE = 0.1682$. There was no benefit or cost when the irrelevant stimulus was the same as the previous relevant stimulus, $F(1, 23) <1$. There was a significant reduction on success rates when the current relevant stimulus was the same as the previous irrelevant stimulus (a negative priming effect of .057), $F(1, 23) = 5.48$, $p < .05$, $MSE = 0.0211$. A similar cost (.061) was found when the entire stimulus ensemble was repeated, $F(1, 23) = 4.43$, $p < .05$, $MSE = 0.0300$, which might primarily reflect negative priming. Note that none of the repetition conditions interacted significantly with S-R pair, $F\text{s}(2, 46) < 1$. Thus, it would again have made little difference if the main analyses had excluded trials with these forms of stimulus or response repetition.

Discussion

Experiment 2 was designed to evaluate whether the switch cost reflects utilization failure or a structural limitation. Despite using a novel time-deadline procedure to provide strong incentives for optimal performance, the switch cost was still substantial (.084 overall). This finding strengthens the conclusion that switch costs reflect the inherent difficulty of task preparation rather than low effort (neglecting to prepare).

As in Experiment 1, the switch cost on success rate was small and non-significant for the first S-R pair (.027), suggesting that participants almost always prepared for this particular S-R pair on task-switch trials. In contrast, the switch cost was substantial for the third S-R pair (.148), suggesting that participants often did not complete preparation for this particular S-R pair on task-switch trials. The switch cost for the second S-R pair (.076) fell somewhere between these two extremes, suggesting that task-switch trials consisted of a mixture of prepared and unprepared states for the second S-R pair. This overall pattern of results is consistent with the partial mapping preparation hypothesis, which assumes that switch costs stem from incomplete S-R mapping preparation for a task switch. In contrast, this pattern of results is problematic for FTE theory, which assumes that switch costs stem from the proportion of task-switch trials on which participants completely fail to engage in advance preparation (which should produce costs for all three S-R pairs).

Experiment 3

The goal of Experiment 3 was to further increase the likelihood that participants would engage in advance preparation. To do so we made several modifications to the design of Experiment 2 (while still using time deadlines). First, to further assist participants in tracking task sequence, we presented task labels on the screen in all blocks rather than just the first 3 practice blocks (as in Experiment 2). Second, the frame for the two boxes corresponding to the color task was presented in purple while the frame for the two boxes corresponding to the shape task remained white. Although purple is not a task-relevant color (the stimuli were red, green, and blue), it may provide a subtle reminder of what task needs to be performed on the upcoming trial. Third, participants were instructed to write down their success rates after each block. This change was intended to make participants more aware of their performance levels and more accountable (knowing that they will have to present this sheet to the experimenter at the end of the experiment). Furthermore, the report sheet also instructed participants to keep improving their performance and to work harder if their success rate fell below .65.

Experiment 3 also addressed two concerns regarding the design of Experiment 2. First, Experiment 2 used different time deadlines for the color and shape tasks, which were adjusted before each task switch (depending on participants' performance). It is therefore possible that the switch costs observed in Experiment 2 were caused not only by the switch in task but also by the switch in the deadline time. To solve this problem in Experiment 3, we applied the same time deadline to the color and shape tasks and made no adjustments within a block.

A second concern was that, for the color and shape tasks, the same stimuli were mapped to each response key for all participants (e.g., red was always mapped to the left response key). This S-R arrangement raises the question of whether the absence of switch costs for the first S-R pair was due to its position or to the specific stimulus itself. To avoid this possible confound, we varied the S-R mapping for each task across participants.

Method

Participants. There were 36 participants in this experiment, drawn from the same participant pool as in Experiments 1 and 2. None had participated in those experiments.

Apparatus, stimuli, and procedure. The apparatus, stimuli, and procedure were the same as in Experiment 2, except as noted. The same time deadline was applied to both the color and shape tasks and did not change within a block. The initial deadline was 3 seconds, which then declined steadily over the two practice blocks until it reached a terminal value of 1 second. Following these two practice blocks, the deadline was held constant across all trials within a block. However, following each block we adjusted the deadline using a tracking procedure designed to achieve an average success rate of .8 for task-repetition trials[8].

To help participants track the task sequence, the frame outlining the color-task boxes was purple. Furthermore, the task labels "color" and "shape" remained on the screen throughout the whole experiment. The stimuli for the color and shape tasks were assigned to each response key across participants using a Latin Square design. Consequently, each stimulus was assigned to each response key equally often across participants. Also, participants were instructed to write down their success rates for the color and shape tasks after each block.

Results

Data analyses were identical to those of Experiment 2. Similar to Experiment 2, participants responded before the deadline on 93% of trials. Responses made before the deadline were 80% correct. The mean deadline was 699 ms, ranging from 524 ms to 1,001 ms across participants.

Success Rates. Table 3 shows the success rates (in proportion) for the task-repetition and task-switch trials. There was a significant switch cost of .049, $F(1, 35) = 12.25$, $p < .01$, $MSE = 0.0103$. The between-experiment comparison showed that this switch cost was not significantly different from that observed in Experiment 2 (.084), $F(1, 58) = 3.12$, $p > .05$, $MSE = 0.0089$. There was a main effect of S-R pair, $F(2, 70) = 9.33$, $p < .001$, $MSE = 0.0145$. Success rate decreased from the first S-R pair to the third S-R pair.

As in Experiment 2, switch costs varied significantly across S-R pairs, $F(2, 70) = 7.48$, $p < .01$, $MSE = 0.0055$. In the follow-up analyses, the simple main effect of S-R pair was significant for the task-switch trials, $F(2, 70) = 14.89$, $p < .001$, $MSE = 0.0105$, but not for the task-repetition trials, $F(2, 70) = 2.03$, $p = .14$, $MSE = 0.0094$. For the task-switch trials, the success rate decreased from the first S-R pair to the third S-R pair (see Table 3). As in Experiment 2, the switch cost was not significant for the first S-R pair (.007), $F(1, 35) < 1$, but was significant for both the second S-R pair (.049), $F(1, 35) = 5.35$, $p < .05$, $MSE = 0.0074$, and the third S-R pair (.091), $F(1, 35) = 25.44$, $p < .001$, $MSE = 0.0067$.

RT for Success Trials. Data analyses on RT were conducted for success trials only. The overall switch cost on RT (5 ms) was not significant, $F(1, 35) = 3.35$, $p = .08$, $MSE = 382$; mean RT was 526 ms for the task-repetition trials and 531 ms for the task-switch trials. There was a main effect of

S-R pair, *F*(2, 70) = 11.46, *p* < .001, *MSE* = 999; RT was 516 ms, 529 ms, and 541 ms for the first, second, and third S-R pairs, respectively. The switch cost varied significantly across S-R pairs, *F*(2, 70) = 9.13, *p* < .001, *MSE* = 512. Similar to the pattern observed in the success rate data, the switch costs on RT were –13, 13, and 15 ms for the first, second, and third S-R pairs. This result indicates that the switch cost on success rate was not due to a speed-accuracy tradeoff.

Stimulus/Response Repetition. We conducted a separate set of data analyses to evaluate the effect of different types of repetitions on success rates for both task-repetition and task-switch trials. Table 4 shows the success rates as a function of different types of repetitions from one trial to the next. Repeating the irrelevant stimulus, without repeating the relevant stimulus, had little effect on success rates relative to the no-repetition condition, *F*(1, 35) < 1. However, repeating the relevant stimulus did substantially increase success rates both when the irrelevant stimulus was not repeated (an increase of .053), *F*(1, 35) = 6.11, *p* < .05, *MSE* = 0.0249, and when the irrelevant stimulus was also repeated (an increase of .132), *F*(1, 35) = 60.75, *p* < .001, *MSE* = 0.0155. These results again are consistent with the hypothesis that participants were more likely to utilize response-selection short-cut strategy when relevant stimulus was repeated. As in Experiments 1 and 2, the effect of repeating either the relevant stimulus itself or the entire stimulus ensemble did not interact with S-R pair, *F*s(2, 70) ≤ 3.07, *p*s > .05. These results suggest our main findings – much smaller switch costs for first S-R pair than for the second and third S-R pairs – would have been obtained even if we had included relevant stimulus repetitions.

For the task-switch trials, the pattern of results was similar to that of Experiment 2. Repeating the response had no effect on success rate relative to the no-repetition condition, *F*(1, 35) = 1.33, *p* = .2562, *MSE* = 0.0255. There was also no benefit or cost when the irrelevant stimulus was the same as the previous relevant stimulus, *F*(1, 35) <1. Success rate was significantly lower when the current relevant stimulus was the same as the previous irrelevant stimulus (a negative priming effect of .054), *F*(1, 35) = 9.68, *p* < .01, *MSE* = 0.0345. Repeating the entire stimulus ensemble slightly reduced success rates (by .034) relative to the no-repetition condition, *F*(1, 35) = 7.58, *p* < .01, *MSE* = 0.0242. None of the repetition conditions interacted significantly with S-R pair, *F*s(2, 70) < 1. As with the previous experiments, the main findings do not depend on whether subsets of data with stimulus or response repetition were included or excluded.

Discussion

In Experiment 3, we made several modifications to further increase the likelihood that participants would engage in advance preparation. Nevertheless, a switch cost was still evident (.049); the switch cost was especially large for the third S-R pair (.091). These results strength the case that participants are incapable, through deliberate preparation alone, of achieving a preparatory state equivalent to that obtained through task repetition.

As in Experiment 2, the switch cost on success rate depended strongly on the S-R pair (see Table 3). The virtual absence of switch cost for the first S-R pair (.007) supports the hypothesis that participants almost always prepared for this pair on task-switch trials. In contrast, the large and significant switch cost for the third S-R pair (.091) indicates that participants often did not fully prepare for this particular pair on task-switch trials. The switch cost for the second S-R pair (.049) fell somewhere between these two extremes, perhaps reflecting a mixture of prepared and unprepared states. Overall, these results further support the partial mapping preparation hypothesis that task-switch trials consist of a mixture of trials with prepared and unprepared S-R pairs. The FTE theory – that task-switch trials consist of a mixture of prepared and unprepared trials – cannot easily account for these results.

General Discussion

The present study examined two key questions about limitations on advance preparation for a task switch. First, we investigated whether advance preparation for an upcoming task-switch trial is all-or-none (the FTE theory) or partial (the partial mapping preparation hypothesis). Second, we investigated whether limitations on advance preparation reflect utilization failure (the FTE theory) or a structural limitation (the partial mapping preparation hypothesis).

Advance Preparation Is Partial, Not All-or-None

As noted earlier, previous task-switching studies typically used only two types of S-R pairs for each task (e.g., even or odd for the parity task). In such cases, there is no obvious reason why all participants would consistently prepare for the same S-R pair (e.g., odd) across trials. To induce a consistent preparatory priority across S-R pairs, we created tasks with three stimuli mapped onto 3 responses (arranged from left to right). We reasoned that participants would prepare the S-R pairs in the left-to-right order of English reading, reinforced by using the same left-to-right presentation order in the instructions. Accordingly, we hypothesized a strong preparation bias in favor of the first (left) S-R pair and against the third (right) S-R pair.

In all three experiments, we found very little switch cost for the first S-R pair. In Experiment 1, the switch cost for this pair was -20 ms. Using the deadline procedure, the switch costs on success rate for the first S-R pair were minimal (.027 in Experiment 2 and .007 in Experiment 3). The CDF analyses for Experiment 1 revealed that the RT distributions for the task-switch and task-repetition trials were nearly identical at all percentiles for the first S-R pair (see Figure 4, top panel), supporting the hypothesis that the first S-R pair was virtually always prepared on task-switch trials. FTE theory provides no obvious explanation for these findings. If participants completely failed to engage in advance preparation on a proportion of task-switch trials, then RT should have been slow regardless of what stimulus happened to appear. Hence, switch costs should have been observed even for the first S-R pair.

Although switch costs were absent for the first S-R pair, they were substantial for the second and third S-R pairs. In Experiment 1, using the standard RT procedure, the switch cost was 107 for the second S-R pair and 117 ms for the third S-R pair. In Experiments 2 and 3, using the deadline procedure, the switch costs were .076 and .049 for the second S-R pair and were .148 and .091 for the third S-R pair. These data suggest that the second and third S-R pairs were often unprepared. In fact, the CDF analyses in Experiment 1 suggest that the third S-R pair was rarely fully prepared – the CDFs for the task-switch and task-repetition trials were separated substantially at all percentiles (see Figure 4, bottom panel). Again, FTE theory provides no obvious explanation for these findings. If participants were fully prepared on a proportion of task-switch trials, then the RT distributions for the task-repetition and task-switch trials should have been the same at small percentiles (the short-RT tail of the distributions) for all S-R pairs.

The consistent finding that the first S-R pair was virtually always fully prepared, while the third S-R pair was not, indicates that participants were often in a state of partial preparation. We conclude that participants can prepare completely for a subset of the S-R mappings for an upcoming task switch (even in the absence of the stimulus), but typically cannot prepare for all S-R pairs. Put more abstractly, it is difficult for top-down control mechanisms to activate an entire "task set" as a coherent entity.

Fundamental Inability to Complete Advance Preparation

De Jong (2000) proposed that participants are capable, through deliberate advance preparation alone, of achieving a preparatory state equivalent to that obtained by recently performing a task. According to this theory, switch costs occur only because participants occasionally fail to engage their competence. The primary support for this strategic switch-cost hypothesis was the apparent evidence for all-or-none preparation: if participants can fully prepare on some trials, then it suggests that they have the competence to prepare on every trial. By refuting the evidence for all-or-none preparation across trials, the present study has undermined the positive case for the strategic switch-cost hypothesis.

Results from Experiments 2 and 3 provide a different line of evidence against the strategic switch-cost hypothesis. Because the standard RT procedure used in task-switching studies does not provide strong incentives for maximal task preparation, we devised a new procedure to do so using time deadlines. In this procedure, success requires a correct response to a moving stimulus before it reaches the center of the screen. If participants fail because their response was too late, they are given clear and immediate feedback. Consequently, a lack of advance preparation would produce an increased proportion of failure trials (either missing the time deadline or making an error) with immediate negative feedback. Because participants do not want to fail, our time-deadline procedure should have discouraged any behavior that would produce slower responding, including failures to prepare for a task switch.

Even with our strong incentives for advance preparation, substantial switch costs were observed on success rates in Experiment 2 (.084 overall and .148 for the third S-R pair). In Experiment 3 we provided even stronger environmental support for engaging in advance task preparation (using both color and word cues as reminders of what tasks would be forthcoming), but we still observed substantial switch costs (.049 overall and .091 for the third S-R pair). Thus, we found no evidence that switch costs could be eliminated by providing stronger incentives and/or stronger environmental task cues. This strengthens the case that switch costs reflect a structural obstacle to complete preparation rather than occasional failures to utilize a latent capability to prepare completely for a task switch.

Do Complete Failures of Advance Task Preparation Ever Occur?

Because we observed little or no switch cost whenever the presented stimulus belonged to the first S-R pair, the data suggest that at least this S-R pair was prepared on virtually every trial. Therefore, complete preparation failures (for the entire S-R mapping) appear to be rare under the conditions we studied. Nevertheless, it is possible that complete failures do occur under other conditions. Rogers and Monsell (1995), for example, reported evidence that participants generally fail to engage in advance preparation when long and short RSIs are intermixed within blocks. They concluded that, in the mixed-RSI condition, "participants may have been reluctant to use the R-S interval to reconfigure in anticipation of a task change when there was a significant probability that the reconfiguration process would be interrupted at any moment by the arrival of the next stimulus" (p. 218). By using only long RSIs, our experiment should have avoided such a problem, and yet task-switch preparation remained far from complete.

Degree of Advance Preparation for S-R Pairs

Previous researchers have pointed out several different obstacles that people might face on task-switch trials. For instance, the task set might not be completely reconfigured (e.g., Goshke, 2000; Rogers & Monsell, 1995; Rubinstein et al., 2001), the relevant task set might still be inhibited from an earlier trial (e.g., Mayr & Keele, 2000), and the irrelevant task set might still be activated from the previous trial (e.g., Allport et al., 1994). Given all these obstacles, it is surprising that there could ever be a dual-affordance condition where task-switch performance exactly equals task-repetition performance. Nevertheless, we found essentially no switch cost for the first S-R pair in all three of our experiments. This result, even if only approximately true, would seem to demand some modification of any hypothesis predicting universal obstacles to task preparation.

For the third S-R pair, we found substantial switch costs in all three experiments. Furthermore, the CDF data from Experiment 1 support the hypothesis that this S-R pair was rarely fully prepared. However, these data do not necessarily indicate that this S-R pair was completely unprepared. At a minimum, some "goal-shifting" might take place during the long RSI, as suggested by traditional task-switching theories (e.g., Goschke, 2000; Rubinstein et al., 2001). That is, participants might remind themselves of which task needs to be performed on the upcoming trial. Furthermore, it is difficult to rule out the possibility that, sometime during the RSI, participants retrieve all three S-R pairs but are then unable to maintain preparation for the full set of S -R pairs. In such a case, the second and third S-R pairs might generally be weakly prepared, but not completely unprepared.

Top-Down and Bottom-Up Forms of Preparation

Our results point to a difference in the effectiveness of deliberate task preparation (i.e., the "top-down" preparation for an upcoming task switch) with the effectiveness of actually doing a task (i.e., the "bottom-up" consequences of task repetition). We found that responses on task-repetition trials are, overall, faster and more accurate than responses on task-switch trials, even excluding trials in which the same S-R pair repeated. Thus, the performance of one S-R pair automatically leads to a heightened state of preparation not just for that S-R pair just performed, but for all the other S-R pairs as well. In contrast we found that deliberate "top-down" advance preparation for a task switch extended to only one, or occasionally two, S-R pairs.

The conclusion that bottom-up and top-down forms of task preparation are qualitatively different was arrived at earlier by Ruthruff, Remington, and Johnston (2001) based on a completely different line of argument. They found that the effects of expectancy and repetition on RT were almost exactly additive, supporting the conclusion that these factors influence separate stages of processing (following the additive factors logic of Sternberg, 1969).

To explain the current results, we propose a more specific hypothesis to account for how "top-down" task preparation works. Our hypothesis is that preparation amounts to rehearsal in a very short-term, limited-capacity buffer, possibly the same as "immediate memory" or "primary memory" in the memory literature (e.g., Waugh & Norman, 1965). Immediate or primary memory is hypothesized to be something like "the current contents of consciousness." As such, one would not think people could usually keep more than one S-R mapping actively in mind at one time. If the mental code produced by an incoming stimulus (e.g., "red") matches the code for the "S" of an S-R pair stored in immediate memory, then one can quickly retrieve the stored abstract "R" code (e.g., "first finger") and use it to generate a response. If a different stimulus appears, then there is no match in primary memory and processing will follow other (slower) pathways. This hypothesis is consistent with recent findings in the memory literature that "focal attention has a very limited capacity, perhaps being restricted to one item or semantically related unit" (e.g., McElree, 2001, p. 830; see also Cowan, 2001).

Although we lack an equivalently specific model of how actual task performance produces a preparatory state, we have already argued that it is unlikely to be the same state. Somehow, performing a task has a deeper effect, spreading activation to all the S-R pairs for that task in some longer-term memory. One more specific possibility for our experiments is that determining the proper response for one stimulus involves determining its ordinal position relative to the two other stimuli; if the ordinal positions of the three stimuli for that task are then stored in some form of cache memory, they could be used to rapidly select the response on a subsequent task repetition.

In addition to this bottom-up form of preparation, task-repetition trials might also benefit from some degree of top-down preparation. For instance, it is possible that task repetitions benefit both from the bottom-up activation of all S-R pairs as well as from the top-down preparation for a specific S-R pair. On the other hand, it has previously been proposed that task repetitions involve relatively little top-down preparation and that, in fact, efforts to impose top-down control actually worsen task repetition performance (Lien & Ruthruff, 2003).

How Does Incomplete Advance Preparation Result in a Switch Cost?

Thus far we have concluded that failure to prepare for all S-R pairs of a new task is the result of a structural limitation rather than a utilization failure. In this section, we move on to the question of how incomplete S-R mapping preparation impairs task-switch performance.

The new deadline procedure used in Experiments 2 and 3 produces a substantial proportion of erroneous responses, and the distribution of those errors on switch trials might provide new clues about what is happening. Three different categories of failure can occur as a result of lack of preparation. First, participants might fail to respond before the deadline – a late response. Or participants might respond in time but choose either of two incorrect responses, one of which was associated with the irrelevant task and one of which was not (the neutral response). Table 5 shows the proportion of task-switch and task-repetition trials with each of these three error types (late response, neutral response, and irrelevant-task response) for each S-R pair in Experiments 2 and 3.

Because our aim is to determine how a lack of advance preparation causes switch costs, we will concentrate on the second and third S-R pairs (the conditions for which advance preparation should be poorest). Table 5 shows the increase in the proportion of each error type on task-switch trials relative to task-repetition trials. Late responses (averaged across the second and third S-R pairs) increased by only .026 and .015 in Experiments 2 and 3, respectively. Neutral-response errors also increased by only a small amount (.021 in Experiment 2 and .007 in Experiment 3). In contrast, irrelevant-task response errors (responses to the irrelevant stimulus) increased substantially; the increase was .070 in Experiment 2, $t(24) = -4.69$, $p < .0001$, and .047 in Experiment 3, $t(36) = -5.12$, $p < .0001$. This increase in irrelevant-task response errors was significantly greater than the increase in neutral responses errors in both Experiment 2, $t(24) = -2.54$, $p < .05$, and Experiment 3, $t(36) = -3.45$, $p < .001$.

The prevalence of irrelevant-task responses on task-switch trials suggests that both the relevant-task response and the irrelevant-task responses were highly activated on task-switch trials; under time pressure, participants were forced to guess between them and often made the wrong choice. One possibility is that the irrelevant-task response was activated more (because it was less effectively suppressed) on task-switch trials than on task-repetition trials, making the selection more difficult. Another possibility is that the relevant-task response was activated less on task-switch trials than on task-repetitions, again making the selection more difficult. Both explanations are plausible and consistent with the present results.

Relations to Previous Task-Switching Findings

Our partial mapping preparation hypothesis is not intended to provide a complete account for all known task-switching phenomena. It may be possible, however, to provide a more comprehensive account by combining our hypothesis with some other mechanisms proposed by other researchers. Mayr and Keele (2000) found that when participants had to switch back to the task from which they recently disengaged, performance suffered. This phenomenon has been attributed to "backward inhibition" – participants inhibit a task they just performed if it is likely to interfere on the next trial (see also Hübner et al., 2003; Schuch, & Koch, 2003). Our basic partial mapping preparation hypothesis does not conflict with this proposal. One of the many consequences of top-down preparation of individual S-R pairs (our hypothesis) may be to release that S-R pair from inhibition left over from previous trials on which it was suppressed.

Conclusion

The present study was designed to clarify the nature of advance preparation for a task switch. In particular, we examined whether (a) advance task preparation is partial or all-or-none and (b) whether people are capable or incapable of fully preparing in advance for a task switch. In all three experiments, we found little switch cost for the first S-R pair but substantial switch cost for the third S-R pair. These findings (along with CDF analyses) support the hypothesis that, preparation for an upcoming task switch almost always prepares the first S-R pair but rarely prepares the third S-R pair. In this way, preparation for each task-switch trial was partial rather than all-or-none. In Experiments 2 and 3 we found that switch costs were observed even in a time-deadline procedure that provides much stronger incentives for advance preparation than the traditional RT procedure. This finding provides no support for the hypothesis that lack of task preparation is due to utilization failure rather than a structural limitation.

The present evidence thus makes a strong case against both of the key assumptions of De Jong's FTE theory. We conclude that instead of preparing all of the upcoming task some of the time, people prepare some of the task all the time. Our partial mapping preparation hypothesis proposes that people typically prepare for only a few specific S-R pairs (often only one). Whether a switch cost is incurred or not, depends on whether the presented stimulus happens to be the one prepared for or not. As with a lottery, the results are "all-or-none", but the preparation is partial on all trials.

References

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

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

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology, 13,* 90-102.

Bertelson, P. (1965). Serial choice reaction-time as a function of response versus signal-and-response repetition. *Nature, 206,* 217-218.

Campbell, K. C., & Proctor, R. W. (1993). Repetition effects with categorizable stimulus and response sets. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1345-1362.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87-114.

De Jong, R. (2000). An intention-activation account of residual switch costs. In S. Monsell, & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (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 situations. *Acta Psychologica, 101,* 379-394.

Fox, E. (1995). Negative priming from ignored distractors in visual selection: A review. *Psychonomic Bulletin & Review, 2,* 145-173.

Gilbert, S. J., & Shallice, T. (2002). Task-switching: A PDP model. *Cognitive Psychology, 44*, 297-337.

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

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task set switching. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 331-355). Cambridge, MA: MIT Press.

Hübner, R., Dreisbach, G., Haider, H., & Kluwe, R. H. (2003). Backward inhibition as a means of sequential task-set control: Evidence for reduction of task competition. *Journal of Experimental Psychology: Learning Memory, and Cognition, 29,* 289-297.

Hübner, R., Futterer, T., & Steinhauser, M. (2001). On attentional control as a source of residual shift costs:

Evidence from two-component task shifts. *Journal of Experimental Psychology: Learning Memory, and Cognition, 27,* 640-653.

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

Lien, M.-C., & Ruthruff, E. (2004). Task switching in a hierarchical task structure: Evidence for the fragility of the task repetition benefit. *Journal of Experimental Psychology: Human Learning, Memory, and Cognition, 30,* 697-713.

Lien, M.-C., Schweickert, R., & Proctor, R. W. (2003). Task-switching and response correspondence in the psychological refractory period paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 692-712.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108,* 393-434.

Los, S. A. (1996). On the origin of mixing costs: Exploring information processing in pure and mixed blocks of trials. *Acta Psychologica, 94,* 145-188.

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.

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 817-835.

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: Control of cognitive processes* (pp. 377-399). Cambridge, MA: MIT Press.

Monsell, S., & Driver, J. (Eds.). (2000). *Control of cognitive processes: Attention and performance XVIII.* Cambridge, MA: MIT Press.

Monsell, S., Yeung, N., & Azuma, R. (2000). Reconfiguration of task-set: Is it easier to switch to the weaker task? *Psychological Research, 63,* 250-264.

Nieuwenhuis, S., & Monsell, S. (2002). Residual costs in task switching: Testing the failure-to-engage hypothesis. *Psychonomic Bulletin & Review, 9,* 86-92.

Pashler, H., & Baylis, G. (1991a). Procedural learning: 1. Locus of practice effects in speeded choice tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 20-32.

Pashler, H., & Baylis, G. (1991b). Procedural learning: 2. Intertrial repetition effects in speeded-choice tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 33-48.

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

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 763-797.

Ruthruff, E. (1996). A test of the deadline model for speed-accuracy tradeoffs. *Perception & Psychophysics*, 58, 56-64.

Ruthruff, E., Johnston, J. C., & Remington, R. W. (2003). Success and failure in promoting parallel central processing with incentives. Paper presented at the 44th annual meeting of the Psychonomic Society, Vancouver, B.C., Canada.

Ruthruff, E., Remington, R. W., & Johnston, J. C. (2001). Switching between simple cognitive tasks: The interaction of top-down and bottom-up factors. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 1404-1419.

Schuch, S., & Koch, I. (2003). The role of response selection for inhibition of task sets in task shifting. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 92-105.

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.

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89,* 669-679.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30,* 220-231.

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72,* 89-104.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41,* 67-85.

Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 455-469.

Authors Note

Mei-Ching Lien, Department of Psychology, Oregon State University. Eric Ruthruff, Roger W. Remington, and James C. Johnston, National Aeronautics and Space Administration (NASA) Ames Research Center, Moffett Field, California.

This research was supported by funding from the National Research Council, NASA Grant NCC 2-1325, and the Human Measures and Performance Project of NASA's Airspace Systems Program. We thank Richard Carlson, Stefanie Schuch, and one anonymous reviewer for their comments on an earlier draft of this article. Correspondence concerning this article should be sent to Mei-Ching Lien at Department of Psychology, Oregon State University, Corvallis, OR 97331-5303. Electronic mail may be sent to mei.lien@oregonstate.edu.

Footnotes

1. In the present study, we used a one-to-one mapping of stimulus attributes to response keys for each task (e.g., red to the left response). For experiments that assign multiple stimulus attributes (e.g., odd digits 1, 3, 7, and 9) to a single response key, people might prepare for the pairing of stimulus attributes to responses or the pairing of stimulus categories (even or odd) to responses. When the stimulus category is pre-learned, however, the latter possibility seems the most likely.
2. This restriction is an indirect consequence of our decision to not allow congruence between the relevant and irrelevant stimuli to occur on a given trial.
3. Although the -20 ms effect for the first S-R pair is not significantly below zero, it is surprising that the true difference should even be zero. It does not seem likely that top-down control mechanisms ever prepare anything with 100% effectiveness on all trials. It is possible that the exclusion of stimulus/response repetitions, while necessary for reasons stated in the text, produces a slight overestimate of the baseline RT for task repetitions relative to task switches (for which there is no corresponding condition where exclusion is needed).
4. Given the conclusion that the second S-R pair is often prepared and the third S-R pair is rarely prepared, it is surprising that overall switch cost for the second S-R pair (107 ms) was only slightly smaller than that for the third S-R pair (117 ms). One possible explanation for this finding is that participants were not always prepared for the third S-R pair on task-repetition trials; this occasional lack of preparation would inflate task repetition RT for the third pair and hence reduce the measured switch cost. This hypothesis is supported by the fact that task repetition RT was 45 ms slower, on average, for the third S-R pair than for the other S-R pairs and that the main difference between CDFs was at the largest percentiles.
5. A more conservative phrasing of this conclusion is that the first S-R pair was prepared equally (though perhaps not always completely) for task switches and task repetitions.
6. We thank the anonymous reviewer for suggesting this possibility.
7. Tracking was based on task-repetition trials only and was performed separately for each task. We increased the deadline when participants failed on a task-repetition trial and decreased the deadline when participants succeeded on two consecutive task-repetition trials for that task. The adjustments to the deadline declined over blocks, based on the assumption that the established deadline would become closer and closer to the desired deadline. Specifically, to increase the deadline, we multiplied the previous deadline by (14+block_number)/(13+block_number). To decrease the deadline, we divided the current deadline by this same quantity.
8. The tracking was based on task repetition performance in the most recent block only. Our goal was to adjust the deadline so that participants would succeed on about 80% of the trials. Because we expected participants to improve slightly in the subsequent block, we established a target of 14 successes out of 18 task-repetition trials (77.7% accuracy) in a block. If participants hit this number exactly, then the deadline remained the same. If the total number of successes exceeded 14, then we decreased the deadline by 4% for each extra correct response (e.g., 18 successes would cause the deadline to decrease by 16%). If the total number of successes was below 14, then we increased the deadline by 3% for each extra failure (with a maximum possible increase of 21%).

Table 1.
Mean Response Times in Milliseconds (Proportion of Errors in Parenthesis) in Experiment 1 as a Function of S-R Pair (Fist, Second, and Third) and Task Transition (Task Repetition and Task Switch). S-R: stimulus-response.

Task Transition	Position of S-R Pair			Average
	First	Second	Third	
Task Repetition	677 (.04)	680 (.03)	723 (.06)	693 (.04)
Task Switch	657 (.05)	787 (.07)	840 (.07)	761 (.06)
Switch Cost	***-20 (.01)***	***107 (.04)***	***117 (.01)***	***68 (.02)***

Table 2.
Mean Response Times in Milliseconds in Experiment 1 as a Function of the Type of Stimulus and Response Repetition and the S-R Pair (Fist, Second, and Third). S-R: stimulus-response; Sr: relevant stimulus; Sir: irrelevant stimulus; R: response.

Trial Type	Position of S-R Pair			Average
	First	Second	Third	
Task Repetition				
No Repetition	683	671	735	696
Sir Repetition	656	698	702	685
Sr/R Repetition	610	632	661	634
Sr/Sir/R Repetition	543	558	602	568
Task Switch				
No Repetition	669	790	825	761
R Repetition	646	808	851	768
Sir was previous Sr	646	731	773	717
Sr was previous Sir	678	778	913	790
Sr was previous Sir, Sir was previous Sr	654	801	842	766

Note: Trials are sorted according to the types of repetition that occurred. Note that different types are possible in the task switch and task repetition conditions. The trial type label specifies all the types of repetition that occurred in that condition. Thus, the types listed are mutually exclusive and exhaustive. Note that these types of repetition did not all occur equally often.

Table 3.
Success Rates in Proportion in Experiments 2 and 3 as a Function of S-R Pair (Fist, Second, and Third) and Task Transition (Task Repetition and Task Switch).

Task Transition	Position of S-R Pair			Average
	First	Second	Third	
Experiment 2				
Task Repetition	.804	.729	.713	.749
Task Switch	.777	.653	.565	.665
Switch Cost	***.027***	***.076***	***.148***	***.084***
Experiment 3				
Task Repetition	.778	.789	.740	.769
Task Switch	.771	.740	.649	.720
Switch Cost	***.007***	***.049***	***.091***	***.049***

Table 4.

Success Rates in Proportion in Experiments 2 and 3 as a Function of the Type of Stimulus and Response Repetition and S-R Pair (Fist, Second, and Third). S-R: stimulus-response; Sr: relevant stimulus; Sir: irrelevant stimulus; R: response.

Trial Type	Position of S-R Pair			Average
	First	Second	Third	
	Experiment 2			
Task Repetition				
No Repetition	.815	.716	.694	.742
Sir Repetition	.786	.752	.759	.766
Sr/R Repetition	.881	.766	.829	.825
Sr/Sir/R Repetition	.937	.892	.824	.884
Task Switch				
No Repetition	.793	.681	.614	.696
R Repetition	.739	.651	.569	.653
Sir was previous Sr	.827	.709	.552	.696
Sr was previous Sir	.756	.638	.524	.639
Sr was previous Sir, Sir was previous Sr	.765	.606	.534	.635
	Experiment 3			
Task Repetition				
No Repetition	.770	.800	.745	.772
Sir Repetition	.788	.769	.747	.768
Sr/R Repetition	.879	.827	.769	.825
Sr/Sir/R Repetition	.926	.904	.882	.904
Task Switch				
No Repetition	.788	.741	.660	.730
R Repetition	.802	.782	.679	.754
Sir was previous Sr	.784	.781	.665	.743
Sr was previous Sir	.726	.685	.618	.676
Sr was previous Sir, Sir was previous Sr	.754	.715	.620	.696

Note: Trials are sorted according to the types of repetition that occurred. Note that different types are possible in the task switch and task repetition conditions. The trial type label specifies all the types of repetition that occurred in that condition. Thus, the types listed are mutually exclusive and exhaustive. Note that these types of repetition did not all occur equally often.

Table 5.

Proportion of Trials with a Late Response, Neutral Response Error, and Irrelevant Task Response Error as a Function of Task Transition (Task Repetition and Task Switch) and S-R Pair (First, Second, and Third) in Experiments 2 and 3.

	Position of S-R Pair			***Average***
Error Type	First	Second	Third	
Experiment 2				
Late Response				
Task Repetition	.042	.083	.066	***.064***
Task Switch	.063	.088	.112	***.088***
Neutral Response				
Task Repetition	.059	.057	.084	***.067***
Task Switch	.054	.070	.113	***.079***
Irrelevant Task Response				
Task Repetition	.086	.130	.135	***.117***
Task Switch	.100	.190	.215	***.169***
Experiment 3				
Late Response				
Task Repetition	.056	.064	.071	***.063***
Task Switch	.066	.074	.091	***.077***
Neutral Response				
Task Repetition	.084	.062	.073	***.073***
Task Switch	.058	.056	.092	***.069***
Irrelevant Task Response				
Task Repetition	.087	.087	.111	***.095***
Task Switch	.102	.127	.165	***.131***

Figure Captions

Figure 1. Cumulative distribution functions for each task transition (task repetition and task switch) and response-stimulus interval (RSI) reported by De Jong (2000; Figure 2). The fit was produced by a mixture model with a probability of .51 that participants fully engage in advance preparation.

Figure 2. An example of the stimulus-response (S-R) mappings used in the present study.

Figure 3. An example of the time course of task presentation in Experiment 1. In this example, the response was correct.

Figure 4. Cumulative distribution functions for each stimulus-response (S-R) pair and task transition (task repetition and task switch) in Experiment 1.

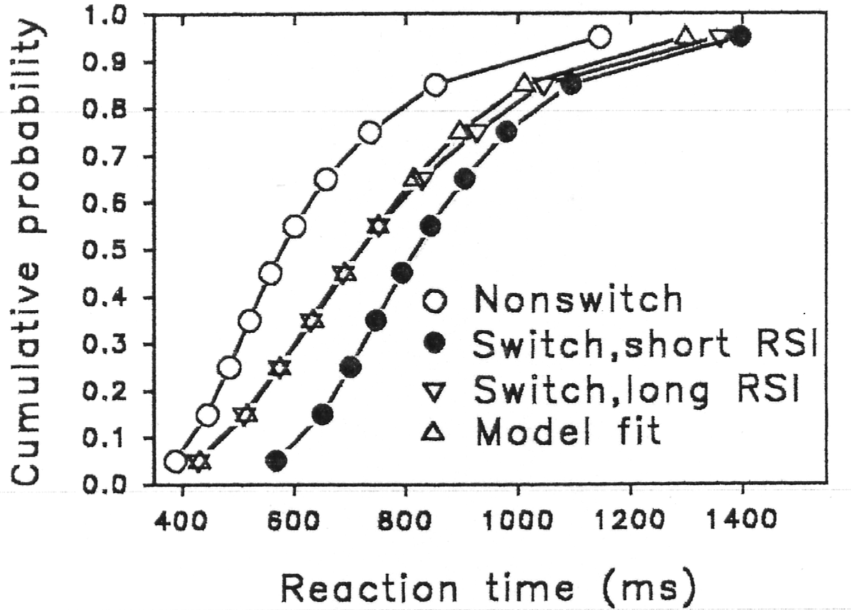

Cumulative probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
400
600
800
1000
1200
1400
Reaction time (ms)
○ Nonswitch
● Switch,short RSI
▽ Switch,long RSI
△ Model fit

Figure 1

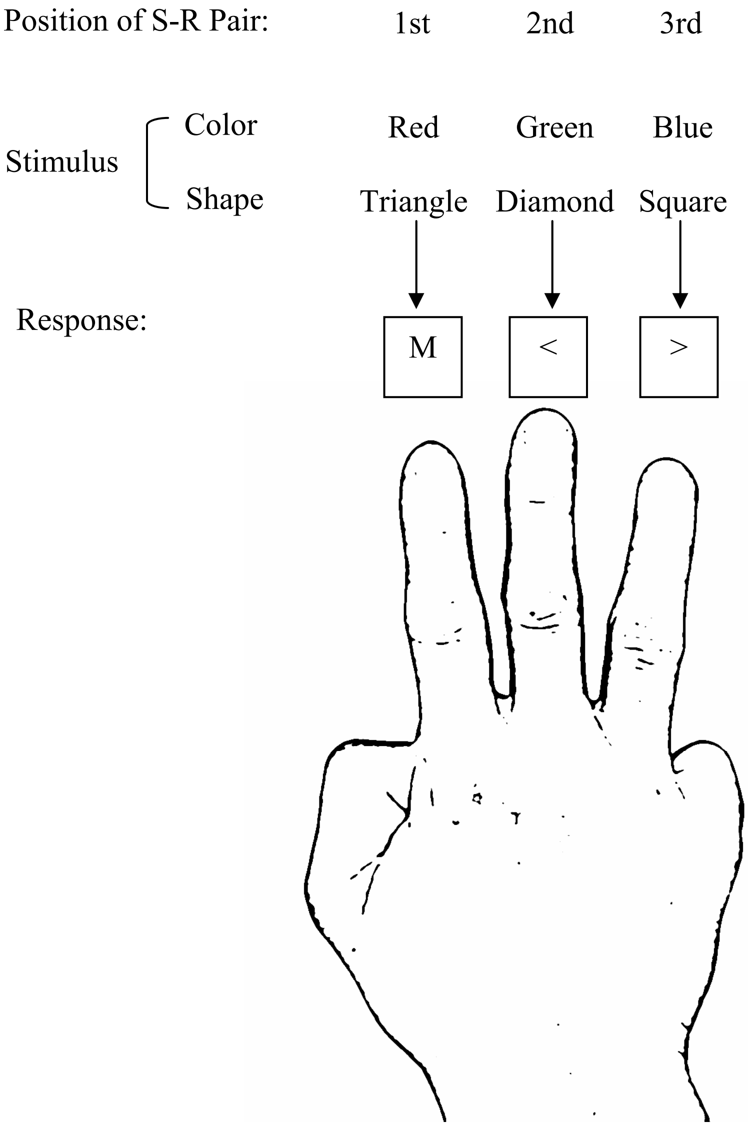
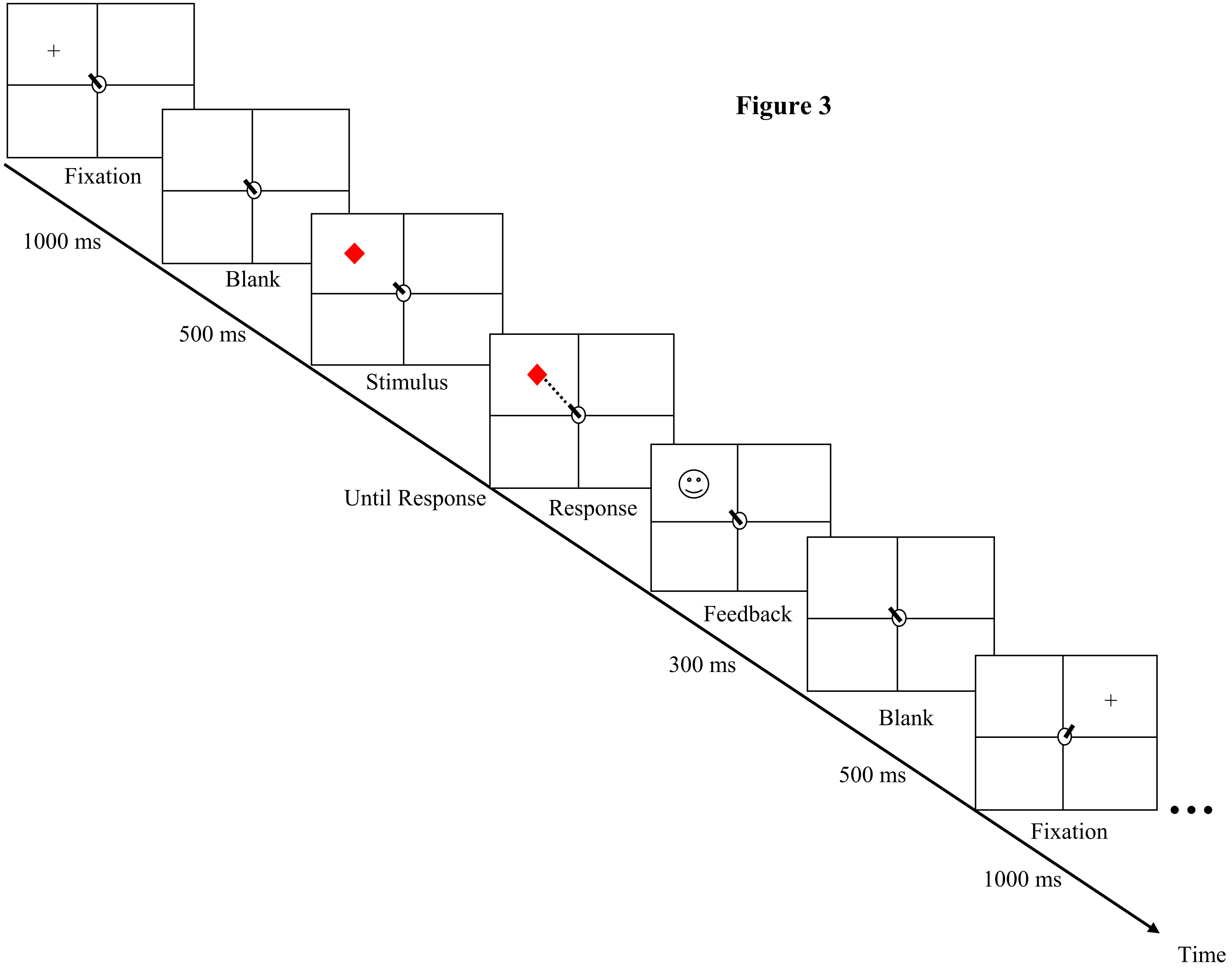
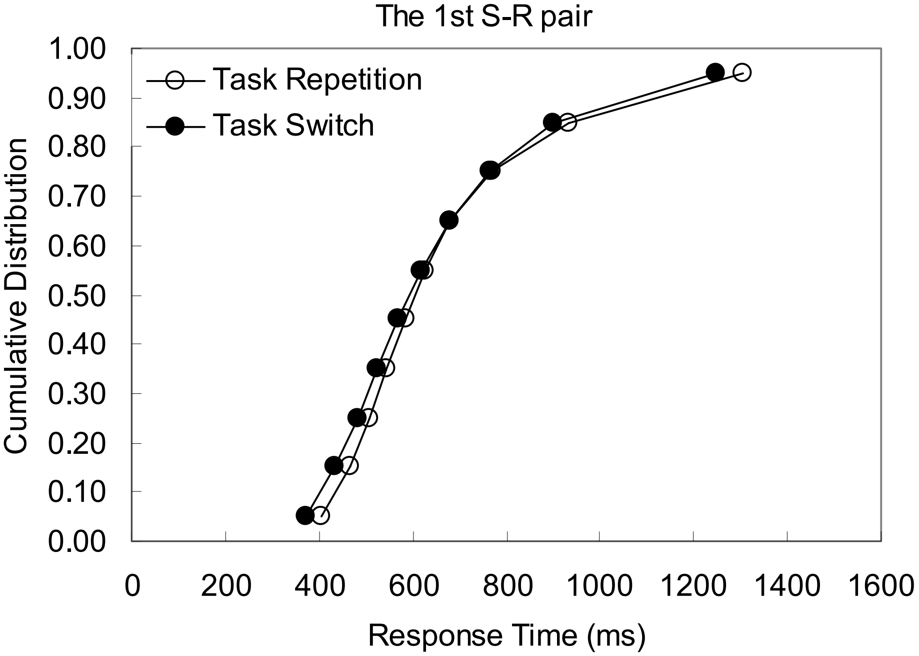
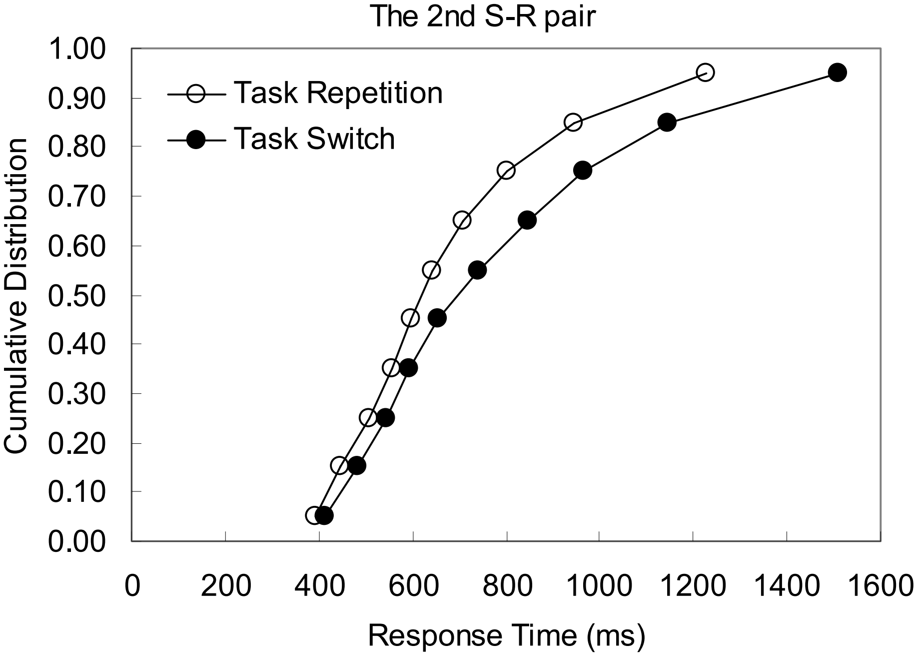

Position of S-R Pair:
1st
2nd
3rd
Stimulus
Color
Shape
Red
Green
Blue
Triangle
Diamond
Square
Response:
M
<
>

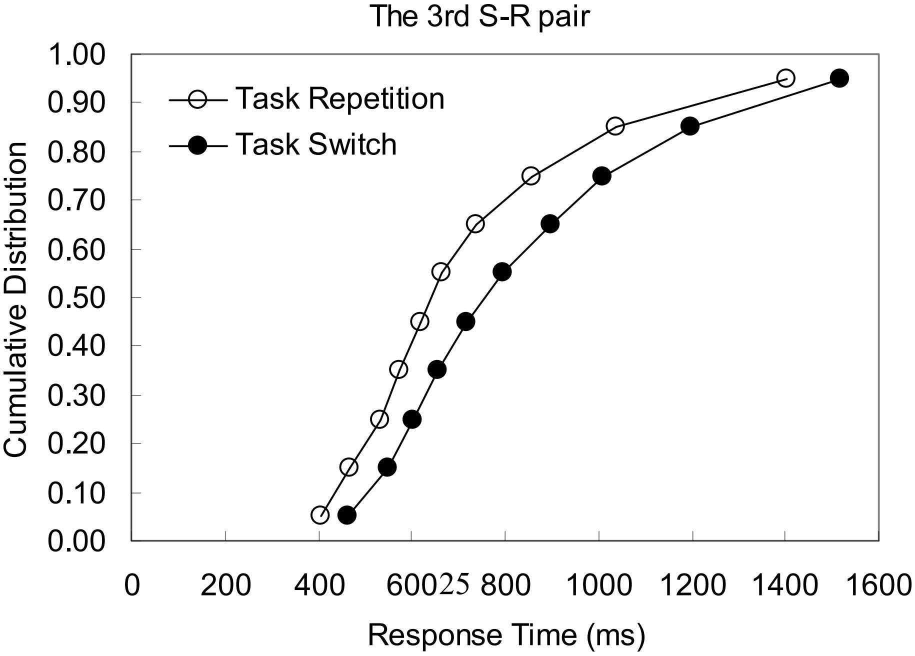
Figure 2

Figure 3


+
Fixation
1000 ms
Blank
500 ms
Stimulus
Until Response
Response
Feedback
300 ms
Blank
500 ms
+
Fixation
1000 ms
• • •
Time


The 1st S-R pair


The 2nd S-R pair


The 3rd S-R pair