Costs of a Predictable Switch Between Simple Cognitive Tasks

Robert D. Rogers and Stephen Monsell
University of Cambridge

In an investigation of task-set reconfiguration, participants switched between 2 tasks on every 2nd trial in 5 experiments and on every 4th trial in a final experiment. The tasks were to classify either the digit member of a pair of characters as even/odd or the letter member as consonant/vowel. As the response–stimulus interval increased up to 0.6 s, the substantial cost to performance of this predictable task-switch fell: Participants could partially reconfigure in advance of the stimulus. However, even with 1.2 s available for preparation, a large asymptotic reaction time (RT) cost remained, but only on the 1st trial of the new task. This is attributed to a component of reconfiguration triggered exogenously, i.e., only by a task-relevant stimulus. That stimuli evoke associated task-sets also explains why RT and switch costs increased when the stimulus included a character associated with the currently irrelevant task.

It has long been understood that cognitive processes require control processes to organize them, but these "executive" mechanisms are—to put it mildly—poorly understood. Failures of control have been identified in patients with damage to the frontal lobes and in the occurrence of certain kinds of action errors in ordinary life. These graphically illustrate the importance of control processes in securing normal performance but have not yet provided a clear picture of their architecture. In the cognitive laboratory, every time we ask participants to perform an experimental task in order to probe processes of interest to us, we engage their control processes as well. But control is rarely the focus of the enquiry—more often it is a nuisance—so that we typically discard data from trials on which the influence of control processes is manifest: that is, early in practice and in recovery from an error. There are certain well-known experimental paradigms, such as the Stroop paradigm and those requiring simultaneous performance of two tasks with similar inputs, processing demands, or memory loads, in which difficulties in control of processing clearly play a major role in limiting rapid and accurate performance. But even in these cases, the role played by control processes in most theoretical accounts remains shad-

owy—a placeholder for an explanation rather than an explanation.

If we are to make progress in understanding control processes, we must directly exercise them and measure their contribution to performance with a vigor similar to that with which researchers have exercised and measured the processes of word recognition, mental rotation, retrieval from memory, and so on. One way of doing so was pioneered by Jersild (1927), who introduced the idea of comparing the time taken to complete a sequence of trials when people did or did not have to switch between different tasks on successive trials. Jersild's paradigm was revived by Spector & Biederman (1976) and has recently been developed further by Allport, Styles, & Hsieh (1994). The work reported in this article belongs in this tradition. Ringing a slight but crucial change on Jersild's paradigm, we measure the costs to performance when participants are required to switch between two tasks afforded by the same input, and we interpret those costs in terms of properties of the process of *task-set reconfiguration*.

Task-Set

Each of the objects and events we encounter in the environment affords a range of possible actions in response to it. The appropriate response varies as a function of the task. In the context of the task of dictation typing, the appropriate response upon hearing the word *fire* is to press the appropriate keys on the keyboard; in the context of the task of target shooting, it may be to pull the trigger; in other contexts, it may be to proceed in an orderly manner to the nearest exit.

To form an effective intention to perform a particular task, regardless of which of the range of task-relevant stimuli will occur, is to adopt a *task-set*. Familiar task-sets, such as naming, can be called up from memory. Novel ones can be specified by instructions or other forms of training. In the cognitive laboratory, we routinely require experimental participants to adopt novel and arbitrary task-sets, addressing them in terms such as these:

Robert D. Rogers and Stephen Monsell, Department of Experimental Psychology, University of Cambridge, Cambridge, England.

Correspondence concerning this article should be addressed to Robert D. Rogers, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB England. Electronic mail may be sent via Internet to rdr10@cam.ac.uk.

On each trial a string of letters will appear in the center of the screen; press this key if it is a word, the other key if it is not; please try to respond as quickly as possible while avoiding errors.

Participants readily follow such instructions, albeit with transient uncertainty about which key is which, how exactly to balance speed and accuracy, and so forth. The same participants may with equal ease be induced to perform a wide range of other tasks given the same stimulus: naming it, classifying it as a noun or a verb, naming its color, counting its syllables, pointing to it, and so on.

To perform one of these simple reactive tasks, the participant must somehow chain together and configure an appropriate set of processes linking sensory analysis to motor output. These processes must include categorization of sensory input with respect to a particular attribute or set of attributes; mapping the attribute's value by means of a decision criterion to one of a predetermined set of response categories; and execution of the motor response used to signal that response category. To assemble processing chains for different tasks, we may or may not reuse the same processes, which are sometimes configured in different ways, sometimes not. For example, a word recognition "module" is part of the process chain for both the lexical decision task ("Is this a word?") and a semantic categorization task (such as "Is this an animal name?")—though it may be used in rather different ways (see Monsell, Doyle, & Haggard, 1989). The word-recognition module is not, however, a logically necessary component of the chain of processes required for other tasks such as indicating the position of an object on the screen or naming its color. However, as the occurrence of Stroop interference in the latter case suggests, a module that is logically unnecessary to the task may nevertheless be "automatically" activated by appropriate input.

Hence to adopt a task-set is to select, link, and configure the elements of a chain of processes that will accomplish a task. We acknowledge that it is difficult to define with precision, even in the restricted context of discrete reaction tasks, what constitutes a "task." For example, it seems intuitively clear that naming and lexical decision are different tasks, whereas two versions of a spatial choice reaction task differing only in the spatial separation of the response keys or the stimulus positions are not different tasks but merely parametric variants on one task. Are different mappings between the same stimulus and response sets (as in Dixon & Just, 1986; Duncan, 1979; Shaffer, 1965, 1967) different tasks? For the present, it will not impair our progress to leave this conceptual boundary cloudy.

In everyday life, we constantly configure and reconfigure elements from our repertoire of cognitive operations and representational capacities to perform now one task, now another. How might we seek to isolate and measure task-set reconfiguration processes? Several authors have examined the effects of a precue that indicates the stimulus–response (S–R) mapping according to which the imperative stimulus that follows it is to be processed (e.g., Shaffer, 1965, 1967). For example, Biederman (1973) showed that cuing the participant in advance about which arithmetic operation was

to be performed facilitated performance of simple addition, subtraction, and multiplication tasks. Sudevan and Taylor (1987) had participants classify a digit either as high or low or as even or odd. Performance was facilitated when the task was cued in advance of the stimulus. The general finding that reaction time (RT) is reduced by a "task" precue implies that individuals can benefit from the opportunity to configure task-set in advance. One can proceed to track the dynamics of this preparation (Dixon & Just, 1986; Sudevan & Taylor, 1987). However, one disadvantage of this paradigm is that one must compare performance with a precue to performance with a simultaneous cue; the latter performance must be influenced not only by the lack of opportunity to preconfigure the task-set but also by the time taken to process the simultaneous cue. It is also hard to control for which task the participant has configured in the absence of a valid precue.

In the present experiments, we adopt a different approach and examine the effects of a completely predictable switch between tasks on successive trials. The index of interest is the cost to performance of such a switch relative to the case when the same task is performed on successive trials. The naive expectation is that such a cost will be observed when the response–stimulus (R–S) interval allows too little time to reconfigure task-set in advance but will vanish when a longer interval is allowed for anticipatory reconfiguration between trials, which yields a subtractive measure of reconfiguration time. The reality turns out to be more complex.

Endogenous and Exogenous Control

To some extent, task-set reconfiguration is *endogenously* (internally) driven. That is, we can adopt task-sets at will, in advance of the stimulus, and without foreknowledge of the stimulus identity other than that it will be a member of a specified class. The responsibility for this intentional component of task control is typically attributed to a special executive mechanism—the Will (James, 1890), Controlled Processing (Atkinson & Shiffrin, 1968; Shiffrin & Schneider, 1977), the Central Executive (e.g., Baddeley, 1986), or a Supervisory Attentional System (Norman & Shallice, 1986; Shallice, 1988, 1994)—that is widely supposed to be unitary, resource-limited, functionally distinct from the processes it organizes, and intimately associated with conscious awareness. Although the endogenous component of task-set undoubtedly exists, we refrain from making any such assumptions about it.

In addition to the endogenous component of task-set, there is also ample evidence that stimuli can of themselves activate or evoke in a person a tendency to perform actions (or tasks)[1] habitually associated with them, irrespective of prior intention, and sometimes in conflict with prior intention. We refer to this as *exogenous* control. (The

---

[1] It is a nice question which of these cases should appropriately be described as evocation of an inappropriate *task-set* and which as evocation of an inappropriate *action*. Monsell, Taylor, and Murphy (1995) argued that the former is the correct description of at least a component of Stroop interference.

endogenous–exogenous terminology is borrowed from its application to a similar distinction between two mechanisms for the spatial orienting of attention—e.g., Briand & Klein, 1987). In a clinical setting, striking illustrations of exogenous control are observed following damage to the frontal lobes (Shallice, 1988). In frontal lobe patients, impaired performance on sorting tasks such as the Wisconsin Card Sorting Test is often associated with an increased proportion of perseverative errors, which suggests that the patient is unable to cease responding to a stimulus attribute that has previously been positively reinforced (see Drewe, 1974; Milner, 1963; Nelson, 1976). Even more dramatic is the "utilization behavior," occasionally observed in some frontal lobe patients (Lhermitte, 1983; Shallice, Burgess, Schon, & Baxter, 1989). These patients are unable to inhibit the performance of complete action patterns characteristically associated with everyday objects such as combs, scissors, matches, light-switches, doors, or teabags when they happen to encounter one of these objects in the environment. We also see analogous losses of control in nonpatients in the class of everyday action errors known as "capture errors," in which a person absentmindedly performs an action habitually associated with the context instead of the action intended. The classic example is supplied by William James's anecdote of the man who goes to his bedroom to dress for dinner and finds himself in bed with his pajamas on. (See Reason, 1984, for an extensive review and many other examples).

In these cases, the person seems, either permanently or transiently, to have lost endogenous control over his or her behavior; behavior is instead pathologically dominated by the power of stimuli to evoke task-sets associated with them—that is, it is exogenously controlled. Disexecutive failures of endogenous control after frontal injury are one major motive for the proposal that the frontal lobes are the site of a specialized "executive" mechanism or mechanisms, one of whose functions is to organize domain-specific processes elsewhere in the brain in such a way as to override habitual but currently inappropriate configurations or to establish new configurations. According to Norman and Shallice's (1986) influential formulation, adoption of the appropriate task-set often does not require the intervention of this executive mechanism: A suitable task-set can, in the main, be automatically selected by a process of "contention scheduling," in which competing strengths of association between stimuli available in the environment and stored task-sets (or action schemas), together with mutual competition between the task-sets themselves, cause one task-set to emerge as dominant among those currently activated. The tendency of available stimuli to activate task-sets with which they are associated is determined by such factors as frequency for use, recency, and salience. However, when contention scheduling does not result in the emergence of a winning task-set or when the winning task-set is in some way inappropriate and needs to be overriden, then the supervisory attention system (SAS) intervenes, modulating the activation levels of the competing schemas so that the person's component processes can be reorganized to meet current behavioral objectives. According to this theory, the relation between endogenous and exogenous control of task-set is such that the endogenous SAS modulates, biases, and perhaps even restarts—when appropriate—the competition between task-sets driven by exogenous input. In this formulation, deliberate adoption of a task-set would be seen as an anticipatory biasing of task-set activations.

The difficulty that endogenous control mechanisms have in overcoming exogenous activation of task-sets is seen not only in outright failures of control but also in the slowing of performance when an inappropriate response tendency is activated by the stimulus and must be suppressed. The classic case is Stroop interference in its many variants (see MacLeod, 1991, for a review). Another is crosstalk between concurrent dual tasks observed when the outcome of stimulus analysis in one task happens to map onto a response category for the other task (e.g., Navon & Miller, 1987). In the present article, we probe the interaction between endogenous and exogenous components of task-set control more directly by trying to measure their contribution to performance when participants are forced to switch between tasks.

Experiments on Predictable Task-Switching

Jersild (1927) compared performance in *alternating* trial blocks, in which the person alternated between two tasks (i.e., ABABA ...), with performance in *pure* blocks, in which the individual performed just one of the tasks (i.e., AAAA ... or BBBB ...). The difference in speed of response and accuracy for these two types of block was used to calculate a *switch cost* (or "shift loss" in Jersild's terms), which was taken to be an index of the extra difficulty associated with reconfiguring task-set. Jersild applied this simple method to an impressive variety of tasks. The costs in terms of performance time produced when participants switched between two tasks were sometimes very substantial. For example, when participants switched between adding 6 to a two-digit number and subtracting 3, the switch cost was some 620 ms per item, 34% of the baseline time per item in the pure blocks. In contrast, switching between naming the antonym of an adjective and subtracting 3 from a two-digit number not only failed to yield a switch cost but—according to Jersild—resulted in a small switch benefit.

Jersild's measurement technology was crude: The time to work through a list of stimuli presented on paper was measured with a stopwatch (often by the participants themselves or an assigned partner). Spector and Biederman (1976) took up the paradigm again and were able to replicate the switch benefit for naming antonyms of words and subtracting 3 from two-digit numbers using Jersild's list-presentation technique, which allows free preview of the next stimulus. However, when participants instead had to work through these same stimuli printed individually on cards, there was essentially no significant cost or benefit. Spector and Biederman also observed large switch costs (over 400 ms per item) for switching between adding 3 and subtracting 3 from two-digit numbers. They interpreted these results, together with those of a further experiment in

which the switch cost was reduced when each of the stimuli was accompanied by task-relevant cues (i.e., 34 + 3, 56 − 3, 12 + 3), as evidence that the principal determinant of switch cost is the extent to which the stimulus provides a readily discriminable retrieval cue for the currently appropriate mental operation as the individual switches between the two tasks in the alternating blocks.

Allport and colleagues have recently conducted a major series of experiments using Jersild's (1927) paradigm: A subset is reported by Allport et al. (1994). Two stimulus ensembles were used, each affording two different tasks. One ensemble was of incongruent Stroop color words (e.g., RED printed in green ink); the participant could be required to name either the color or the word. The other ensemble was a set of displays each containing between 1 and 9 tokens of the same digit; the participant could be required to name either the digit ("value") or the number of digits ("group size"); again, these were always incongruent. Allport et al. found that the costs to performance of task-switching were smaller when the participant had to switch task between ensembles (e.g., between color naming and value naming) than when he or she had to switch task within ensembles (e.g., between color naming and word naming). This is in accord with Spector and Biederman's (1976) suggestion that task-switching is easier when the stimulus provides an effective cue for the task required, which it did in the former, but not in the latter, case.

Other results were more surprising. First, the size of the switch cost observed when the participant had to switch between two tasks involving different stimuli (e.g., word naming and digit naming) depended on whether the participant had recently performed the other task afforded by the same stimulus type. For example, switching between word naming and digit naming resulted in dramatic switch costs only for participants who had, earlier in the experimental session, switched between color naming and group-size naming. And switching between these two tasks resulted in virtually no performance costs unless the participants had earlier switched between the word and digit naming tasks. Allport et al. therefore interpreted the occurrence of a switch cost in these paradigms as a form of "pro-active interference" from a recently adopted task-set afforded by the same stimulus type. Their results led them to conclude that this *task set inertia* (TSI)—as they term it—is dissipated only after several minutes of performing other tasks.

The intuitively attractive account of the switch costs observed in Jersild's paradigm is essentially subtractive: The temporal cost of a switch indexes the duration of a control process that accomplishes task-set reconfiguration, and this must be done before the next stimulus can be properly processed (beyond perceptual analysis). Allport et al. (1994) directly tested this account in their Experiment 5. Participants saw just two successive displays on each trial. They were forewarned either to perform the same task in response to both displays or to switch tasks (e.g., name the word on the first display and then the color on the next display). The interval between the response to the first stimulus and the presentation of the second stimulus (R–S interval) was varied. Allport et al. reported no reliable

reduction in switch cost when the R–S interval was increased from 20 to 1,100 ms. If the (substantial) switch cost at a short R–S interval is due to the time taken by a reconfiguration process, why, when the R–S interval is extended to more than a second, could the participant not escape that cost by completing the reconfiguration process before, and in anticipation of, the second stimulus? Allport et al. argued that the switch cost was accounted for not by the duration of an executive reconfiguration process[2] but by poststimulus interference from a (recently activated) competing task-set—task-set inertia. Our own data raise the same question, but we offer a different answer in later discussion.

A New Switching Paradigm: Alternating Runs

Happily unaware of Allport et al.'s (1994) findings, we had meanwhile developed a different way of measuring the cost of task-switching. We did this to avoid two difficulties with Jersild's (1927) estimate of switch cost as the difference between performance in alternating and pure blocks of trials. First, in the alternating blocks the participant must do two things that are not required in the pure blocks: (a) keep two, rather than one, task-sets "active," or available, and (b) reconfigure between them on every trial. It is not clear which of these demands is indexed by the switch cost. Second, a between-blocks comparison is open to the objection that there may be differences in effort, arousal, response criterion, within-task strategy, and so on, between blocks. The alternating blocks are unquestionably perceived by participants as more difficult: It is entirely plausible that this could lead to greater arousal, effort, and so forth. Indeed this might account for the otherwise inexplicable switching benefit that Jersild observed with some task pairs.

A simple solution to both problems is to compare switch and nonswitch trials within a block. One can do this by alternating between runs of trials of predictable length on each task. That is, the participant knows that he or she must switch tasks every nth trial. To help the participant keep track, one accompanies the stimulus by a cue indicating its position in the current run. For all but the last experiment reported here, $n = 2$ and there were two tasks, so that the sequence of trials was AABBAABB, and so on. We could estimate the cost of switching tasks by comparing performance on trials on which the participant had to switch (i.e., AB, BA) with performance on otherwise identical trials on which no switch was required (BB, AA). We refer to this difference in RT as the *time cost*, and to the difference in error percentage as the *error cost*, of the switch. Because the participant must keep both task-sets in a state of adequate

---

readiness, we suggest that this within-block index is a purer measure, and one less vulnerable to confounding with arousal or criterion effects, than the contrast of performance on alternating and pure blocks.

Overview of the Experiments

The six experiments focus on two questions, initially distinct: the effect of varying R–S interval, and the exogenous cuing of task-set and crosstalk between tasks.

The Effect of Varying R–S Interval

Experiment 1 will demonstrate that, for the pair of tasks we use throughout this series, there is a large switch cost when the R–S interval is short. On the assumption that the time cost of a task-switch indexes the duration of an endogenously driven reconfiguration process, we would predict that a sufficiently long R–S interval together with a predictable task-switch would allow the participant sufficient time to reconfigure in anticipation of the task-switch. Experiments 2, 3, and 4 thus vary the R–S interval, the expectation being that as R–S interval increases to some value (i.e., the maximum reconfiguration time) the switch cost will diminish to nothing. Although we do, under some circumstances, detect a decline, the experiments also demonstrate (like Allport et al.'s, 1994, Experiment 5) a large component of switch cost that shows no sign of disappearing even when we increase the R–S interval to more than a second; we call this the *residual switch cost*. In Experiments 5 and 6, we test several accounts of the residual switch cost, including Allport et al.'s (1994) task-set inertia theory.

Exogenous Cuing of Task-Set and Crosstalk Between Tasks

Shifting task-set requires that control mechanisms must somehow decouple or suppress the S–R mapping for the task that was appropriate on the previous trial and will be appropriate again soon but is not now appropriate for the current trial. We define the attribute of the stimulus to which the participant is being asked to respond when performing the current task as the *relevant* attribute and other attributes (to which the participant might respond when performing other tasks) as *irrelevant* attributes. The experiments examined the effect on switch cost of whether an irrelevant attribute was associated with a response in the now-irrelevant task. The hypothesis of exogenous activation of a task-set by stimuli associated with that task suggests that it should be harder to switch task-sets—and in doing so, decouple the now-irrelevant S–R mapping—when the stimulus contains an attribute associated with (a response in) the task being switched away from—an *exogenous cuing effect*.

In addition, in all of these experiments, the stimuli in both tasks were mapped via different criteria to the same response set (keypresses with the left and right index fingers). Hence, to the extent that Task-set A is not successfully suppressed when Task B is required, we may expect to see

crosstalk between the tasks: facilitation when the irrelevant attribute is mapped to the same response as the relevant attribute, interference when it is mapped to the other response. As already noted, such crosstalk sometimes occurs in dual task experiments and is also the source of Stroop interference. We shall show that, though there is a modest crosstalk effect with successive performance of different tasks, it is dominated by an exogenous cuing effect. We shall make use of this outcome in our account of the residual switch cost and suggest that exogenous cuing of task-set may be a factor in other paradigms also.

The Tasks

To investigate these questions, we need the participant to switch between two tasks that meet the following criteria. The two tasks must be mapped to the same response set. Each task must be based on an S–R mapping requiring little task-specific training. The two tasks must be driven by attributes of the same stimulus that are separable—that is, that can be varied independently. The irrelevant attributes must be able to take on neutral values (i.e., values not associated with a response in either task). Finally, the number of attribute combinations must be large: Any combination must be experienced with a low enough frequency to make it unlikely that the participant will, during the course of the experiment, learn an association between each combination of attributes, task cue, and response. Otherwise the participant may, through learning, redefine what the experimenter wishes to treat as two tasks, as one. (We include a check on this point in Experiment 1.)

The two tasks we used were classifying a letter as a consonant or a vowel and classifying a digit as odd or even. Each stimulus consisted of a pair of characters presented side by side. In the letter task, one (and only one) of the characters was a letter; the participant pressed one key if it was a consonant, the other if it was a vowel (see Figure 1a). In the digit task, one (and only one) of the characters was a digit; the participant pressed one key if it was odd, the other if it was even. The relevant character was randomly on the right or the left of the pair, and the two characters were closely adjacent. We assumed that the participant would not be able to avoid processing the irrelevant character to some degree—an assumption borne out by the effects we will report. But we make no assumptions about whether the two characters are perceptually processed to the same "depth," or the degree to which they are processed in parallel, or whether there is any bias to give priority to the left or right character position.

In one condition—the *no-crosstalk* condition—the irrelevant character was always drawn from a third set of non-alphanumeric characters, which we will call the *neutral* set. Because these characters were associated with responses in neither task, each character pair afforded a response only with respect to the appropriate task. In the other condition—the *crosstalk* condition—the irrelevant character was drawn from the neutral set on only one third of the trials. On the other two thirds the character pair contained a letter and a

digit, so that the irrelevant character was one of those associated with the currently inappropriate task. The response (left or right) associated with the irrelevant character was on one third of the trials *congruent* with that dictated by the relevant character, and on one third *incongruent*. Inasmuch as crosstalk occurs, we would expect faster responses to congruent than to incongruent pairs. But note that only two thirds of the stimuli in this condition afforded a response in both tasks, and hence afforded the possibility of crosstalk. By comparing performance on trials where the irrelevant character is neutral with performance on trials where it is non-neutral, we can assess the exogenous task-cuing effect—the degree to which an irrelevant attribute associated with a now-inappropriate task makes that task harder to suppress, irrespective of its particular response value.

Experiment 1

In this experiment we compared the cost of switching, predictably, between the letter and digit tasks in the crosstalk and no-crosstalk conditions, with an interval between each response and the following stimulus of only 150 ms, over 2 days of practice.

Method

Participants. There were 8 paid participants, 1 woman and 7 men, between 21 and 46 years of age.

Procedure. The participants sat with their eyes approximately 40 cm from the screen of a Macintosh SE/30 computer on which was displayed a framework of four square boxes (see Figure 1b). The centers of the four boxes defined four positions at the corners of a 5-cm square. On each trial a character pair was displayed in one such position, as in Figure 1b. On successive trials, the position moved to the next square clockwise. One participant might be told to perform the letter task when the character pair appeared in either of the top two positions and the digit task when the character pair appeared in either of the bottom two positions. Another participant would receive a different assignment of pairs of adjacent positions to tasks. For every participant, successive displays proceeded in a clockwise circuit, and the task changed predictably on every second trial, with the display position providing a reliable external cue to position in the AABB task cycle.

The assignment of pairs of adjacent positions to tasks was counterbalanced over participants so that the direction of the eye movement between trials would not be confounded with task or position in a run.

Each character pair was displayed in an uppercase Courier bold font, which subtended 1.1° horizontally and 0.7° vertically. It remained on the screen until the participant pressed a key or until 5,000 ms had elapsed. There then followed a 150-ms interval before a new character pair was presented at the next position in the clockwise sequence. If the response key had not been released by the time the next character pair was presented, the trial was excluded from the data analysis. If the participant made an incorrect response, a beep sounded for 20 ms and the following R–S interval was extended to 1,500 ms. Experience indicated that without such a recovery period, an error was likely to be followed by others because the participant tended to lose track of which task was now appropriate.

Consonants were sampled randomly from the set G, K, M, and R, vowels from the set A, E, I, and U, even digits from the set 2, 4, 6, and 8, and odd digits from the set 3, 5, 7, and 9. The neutral characters used were #, ?, *, and %. Sequences of stimuli were constructed with the restriction that the same character could not appear on two successive trials.

Between blocks, the word "Ready?" was displayed until the participant pressed the space bar to display the four boxes. The first character pair was presented 2 s later, in a position of which the participant was forewarned. At the end of each block, a feedback display indicated the number of errors and the mean correct RT for that block. The participant was instructed to try to minimize RT while avoiding errors; around 2 or 3 errors per block was suggested as an acceptable level of performance. A rest break was allowed every 4 blocks.

Design. Testing was conducted over 2 days and consisted of several sequences of four blocks each. Each block consisted of 48 trials, the first 12 of which were treated as warm-up trials and excluded from the data analyses. New sequences of character pairs were generated for every participant and day of testing.

On each day, the participant completed two contiguous sequences in the no-crosstalk condition and two contiguous sequences in the crosstalk condition. The order of presentation of these conditions was counterbalanced over participants and days. Before the first experimental sequence of the day in either condition, the participant was given two practice blocks in that condition.

Each four-block sequence of the crosstalk condition contained one experimental trial with every combination of the following variables: task (letter or digit), trial type (switch or nonswitch),
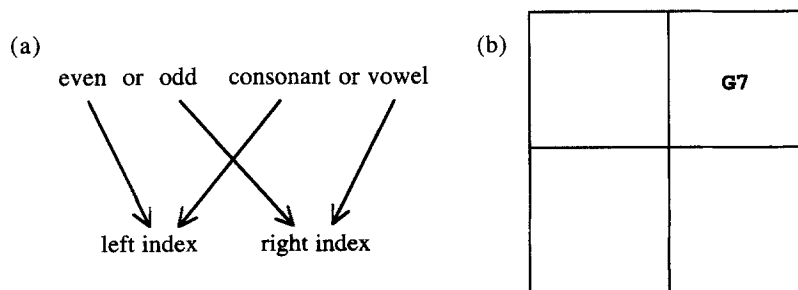
Figure 1. (a) The stimulus–response mappings used by half of the participants in Experiment 1; for the other half, the responses used for the digit task were exchanged; (b) an example of a display as seen by the participant during the crosstalk condition (in Experiments 1–3 and Experiment 6).

response (left or right), response on the previous trial, irrelevant character type (congruent, incongruent, or neutral), and irrelevant character type on the preceding trial. Similarly, each sequence of the no-crosstalk condition contained equal numbers of trials for each combination of task, trial type, response, and preceding response.

Preexperimental training. In other experiments we have found that participants tend to have difficulty switching tasks if they have insufficiently practiced each task on its own, presumably because of having simultaneously to remember both which task is currently appropriate and the new and arbitrary S–R mapping for each task. Although the question of how an S–R mapping is established early in the performance of a new task is itself an important and interesting issue in relation to executive control processes (see, e.g., Shiffrin & Schneider, 1977), the purpose of the present experiments was to investigate the process of switching between well-established task-sets. Hence, on the 1st day of the experiment, before being introduced to the alternating runs procedure described above, the participants were given practice on the letter and the digit tasks, as follows. Character pairs were displayed one at a time in the center of a rectangle 8.5 cm wide and 5.5 cm high. The irrelevant character was always drawn from the neutral set, and the R–S interval was set at 150 ms. The subject alternated between a 24-trial block on the letter task and a 24-trial block on the digit task, for 16 blocks, which yielded 192 trials with each task. Pretests had established that this amount of practice was sufficient to stabilize the speed and accuracy of performance with each task.

During this training and for the first two practice blocks with the alternating runs procedure, a reminder of the response assignments was taped beneath the screen. Four participants were given the response assignments shown in Figure 1a. For the other 4, the keys used for the digit task were exchanged, so that the relationship between the fingers used for "odd" and "consonant" was counterbalanced. A *left* response was made by pressing the c key on the computer keyboard with the left index finger, and a *right* response was made by pressing the m key with the right index finger.

Postexperimental transfer test. As will be seen, performance on both tasks improved with practice. We were concerned with the possibility that although the number of stimulus combinations was reasonably large, participants might nonetheless be gradually learning direct associations between configurations of stimuli, task cues, and individual responses and ceasing to base their response on classification of the stimuli. Any such recasting of the pair of experimenter-defined classification tasks as a single participant-defined S–R mapping task would undermine the rationale for the experiment. To check for this, we followed the main experiment with a transfer test in which the participant continued to switch between the letter and digit tasks, but with a new set of four consonants replacing the original set.

Specifically, when the participant had completed four sequences of blocks on the 2nd day (i.e., two contiguous sequences in each of the no-crosstalk and crosstalk conditions), the original set of consonants was replaced by the set N, P, T, and X. The participant then completed one further sequence (of four blocks) of both the no-crosstalk and crosstalk conditions, in the same order in which these conditions had been performed in the main experiment on the 2nd day. Prior to each sequence, he or she was given two practice blocks with that condition. If performance at the end of the main experiment was based on stimulus-specific responding, we should observe a dramatic regression in performance with the new set of consonants.

Results

In all of the experiments, trials with RTs of less than 100 ms and trials immediately following an error were excluded from the data analyses. Data from trials in which the stimulus display was presented before the participant had released the previous response key were also excluded. In the first experiment, 12.7% of trials were lost in this way. The pattern of data was not changed in any important way if they were included, but the mean RT was slightly increased, which suggests that a few of these responses may have been delayed by lack of motor readiness.

Switch cost. Table 1 shows the RT and error data for the two tasks on the switch and nonswitch trials, for both days of testing. Mean RTs and error proportions[3] were initially examined by repeated measures analysis of variance (ANOVAs) with the following factors: practice (1st day or 2nd day), condition (no crosstalk or crosstalk), task (letter or digit), and trial type (switch or nonswitch). We calculated the switch costs for each task by subtracting the mean RT and error rate on nonswitch trials from the corresponding values on switch trials. Mean RT was reliably greater on switch trials than on nonswitch trials, $F(1, 7) = 66.2, p < .001, MSE = 24,383$; so was the error rate, $F(1, 7) = 37.6, p < .001, MSE = 0.041$. The interaction between task and trial type was not reliable for either measure: $F(1, 7) = 2.7$ for the RTs; $F < 1$ for the error rates. The average time cost incurred by switching tasks was 224 ms, and the error cost was 4.2%.

Switch cost and practice. The extensive practice provided by a second day of testing resulted in a modest reduction in switch cost. The RT cost fell from 262 ms on the 1st day to 186 ms on the 2nd, $F(1, 7) = 8.0, p < .05, MSE = 5,877$, and error cost fell from 5.5% to 3.0%, $F(1, 7) = 4.5, p = .07, MSE = 0.039$. The simple main effect of trial type on the 2nd day was reliable for both RT, $F(1, 7) = 50.5, p < .001, MSE = 11,002$, and error rate, $F(1, 7) = 11.9, p < .05, MSE = 0.028$. However, the baseline performance on the nonswitch trials was also reduced with practice for RTs, $F(1, 7) = 25.6, p < .005, MSE = 5,435$, though not for error rates, $F < 1$.

Task-set cuing and crosstalk. To test the conjecture that adoption of the appropriate task-set is harder when the stimulus includes an attribute associated with the task from which the participant is attempting to switch, we need to look at two contrasts. The first is between the crosstalk condition (in which the irrelevant character was frequently one associated with the other task) and the no-crosstalk condition (in which the irrelevant character was always drawn from the neutral set). Table 2 shows the relevant data, averaged over day of testing. Clearly, the frequent occurrence of irrelevant characters associated with the now-inappropriate task slowed RTs for both switch and non-switch trials, $F(1, 7) = 101.9, p < .001, MSE = 9,464$; though the main effect of condition was unreliable for the errors, $F < 1$. Also, the switch cost in the crosstalk condi-

---

[3] Error proportions were arc-sine transformed; the reported *MSE* is always that of the transformed data.

Table 1

Performance and Switch Costs by Tasks and Days in Experiment 1

Task	1st day		2nd day		Average	
	RT	% error	RT	% error	RT	% error
Letter						
Switch trials	966.0	7.6	810.4	7.7	888.2	7.6
Nonswitch trials	729.2	2.6	630.7	3.0	679.9	2.8
Switch cost	236.8	5.0	179.7	4.7	208.2	4.8
Digit						
Switch trials	986.2	8.0	802.1	4.1	894.1	6.0
Nonswitch trials	697.2	1.9	609.0	2.8	653.1	2.3
Switch cost	289.0	6.1	193.1	1.3	241.0	3.7

Note. RT = mean reaction time in milliseconds.

tion was 128 ms greater than that in the no-crosstalk condition, $F(1, 7) = 15.2, p < .01, MSE = 8.650$; the error cost was not reliably greater, $F(1, 7) = 2.2$. On the 1st day, the switch cost was 168 ms greater in the crosstalk than in the no-crosstalk condition; on the 2nd day, the corresponding figure was 88 ms, but the extra error cost found in the crosstalk condition increased from 0.1% on the 1st day to 2.3% on the 2nd day. This interaction among trial type, condition, and practice was not reliable for RT, $F(1, 7) = 3.0$, but was reliable for error rates, $F(1, 7) = 7.7, p < .05, MSE = 0.012$.

The second relevant comparison is within the crosstalk condition, between trials with congruent, incongruent, and neutral irrelevant characters. This contrast, shown in Figure 2, was examined by ANOVAs with the factors of practice (1st day or 2nd day), task (letter or digit), trial type (switch or nonswitch), and irrelevant character (congruent, incongruent, or neutral). RTs were longer when the irrelevant character was either congruent or incongruent than when it was neutral, $F(2, 14) = 52.0, p < .001, MSE = 11,892$. Newman-Keuls tests showed that latencies on both congruent and incongruent trials were reliably different from those on neutral trials ($p < .01$), but not from each other. The main effect of irrelevant character type was also reliable for

the error rate, $F(2, 14) = 5.4, p < .05, MSE = 0.162$, but with significantly more errors on incongruent trials than on both congruent trials and neutral trials ($p < .05$ in each case).

Analysis of simple effects indicated that the irrelevant character significantly affected RTs on switch trials, $F(2, 14) = 37.0, p < .001, MSE = 15,066$, and on nonswitch trials, $F(2, 14) = 52.5, p < .001, MSE = 2,587$. However, there was also a significantly greater effect of the irrelevant character on switch trials, $F(2, 14) = 13.0, p < .01, MSE = 5,761$. Thus the time cost of a task switch was 309 ms on congruent trials, 350 ms on incongruent trials, but only 216 ms on neutral trials; the first two were reliably larger than the third ($p < .01$ in each case). The error cost of a switch was larger on incongruent trials (7.0%) than on both congruent (4.0%) and neutral (4.2%) trials, but these differences were not reliable, $F < 1$.

Finally, we can compare performance on neutral trials in the crosstalk condition with performance in the no-crosstalk condition (for which all trials were neutral). The main effect of condition was significant for both RT, $F(1, 7) = 41.4, p < .001, MSE = 2,987$, and error rate, $F(1, 7) = 10.4, p < .05, MSE = 0.050$, so that latencies were increased by 62 ms, but error rates reduced by 1.0%, in the crosstalk con-

Table 2

Performance and Switch Costs by Crosstalk Conditions and Tasks in Experiment 1

Condition	Letter		Digit		Average	
	RT	% error	RT	% error	RT	% error
No crosstalk						
Switch trials	776.9	6.9	768.1	6.1	772.5	6.5
Nonswitch trials	619.9	3.1	603.8	2.5	611.8	2.8
Switch cost	157.0	3.8	164.3	3.6	160.6	3.7
Crosstalk						
Switch trials	999.9	8.4	1020.3	6.1	1010.1	7.2
Nonswitch trials	740.2	2.5	702.6	2.2	721.4	2.3
Switch cost	259.7	5.9	317.7	3.9	288.7	4.9

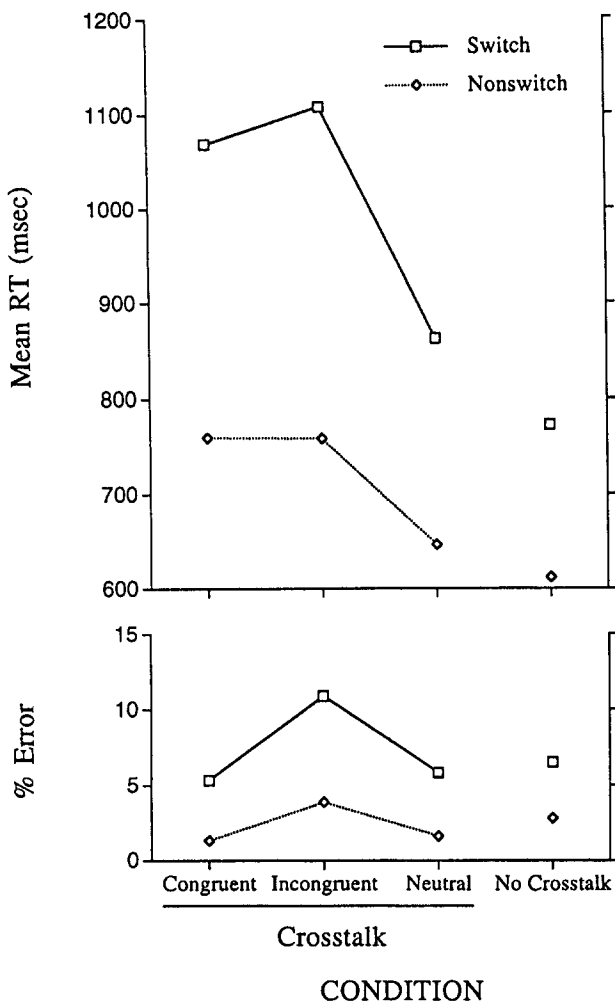*Note.* RT = mean reaction time in milliseconds.

Figure 2. Mean reaction time (RT) and error rate in Experiment 1 on switch and nonswitch trials as a function of the irrelevant character type, averaged over day and task.

dition compared with the no-crosstalk condition. The time cost was also reliably larger (by 55 ms) in the crosstalk condition, $F(1, 7) = 7.2, p < .05, MSE = 3,426$; the error cost was also slightly larger, but not reliably so, $F < 1$. Hence there is some evidence that the slower switching in the crosstalk condition extends to stimuli that are identical in not affording an opportunity for crosstalk.

Postexperimental transfer test. After the main experiment, from which the data so far have been reported, the transfer test, which used a different set of consonants, was given. Participants were given one practice pair of blocks and one test sequence of four blocks, for each of the crosstalk and no-crosstalk conditions, in the same order as used for the earlier part of the 2nd day. Figure 3 shows the data plotted as a function of block sequence: sequences 7 and 8 are the last two sequences from the main experiment, and sequences 9 and 10 are the transfer sequences. The practice sequences preceding them are also included in case any effect of transfer is transient.

There is no obvious discontinuity or worsening in overall performance at the time of transfer. Presumably, any such effect should be maximal for the consonant response in the letter task, because it is the stimuli mapping on to this response that were changed. An ANOVA for the letter task, with sequence (the six levels shown in Figure 3), trial type (switch or nonswitch), decision (consonant or vowel), and order of conditions as the factors showed no main effect of sequence, interaction of sequence with decision, or interaction of sequence with switching for mean RT or errors ($Fs < 1.5$ in all cases). As Figure 3 shows, the only response category to show any signs of a decline in performance in the transfer test was in fact the *odd* response in the digit task, which had not been changed, and for this task the main effect of sequence approached a reliable level for mean RT, $F(5, 30) = 2.39, p = .06, MSE = 5,292$, but not for errors, $F(5, 30) = 1.48$; no interactions involving sequence approached significance for this task.

Discussion

The results of Experiment 1 establish, for two simple character-classification tasks, the basic phenomenon we
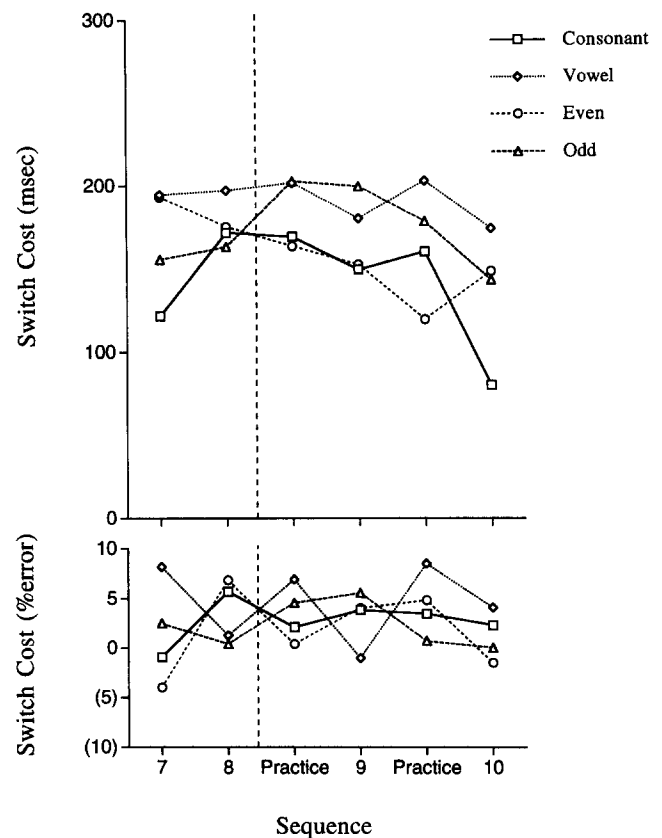

Figure 3. Performance leading up to and during the transfer test at the end of Experiment 1, when a new set of consonants was introduced. Sequences 7 and 8 were the last two four-block sequences of the main experiment. The practice sequences contained only two blocks.

wish to explore in the remaining experiments: a large and reliable cost to both accuracy and speed of performance when a participant must switch predictably but quickly between two tasks. An important advantage of the way we measured this switch cost, as compared with Jersild's technique (Allport et al., 1994; Jersild, 1927; Spector & Biederman, 1976) is that the switch and nonswitch trials are mixed together in the same blocks. Hence the switch cost may be securely attributed to the need to reconfigure task-set, rather than to the need to maintain two task-sets in an available state rather than one, or to differences between blocks in arousal, effort, or response criterion.

It is worthy of note that the switch cost was robust to practice. Not only was the mean switch cost on the 1st day of the experiment based on performance measured after several initial blocks of practice, but the extra practice the participant had received by the 2nd day (about 1,000 trials) diminished the switch cost relatively little. If we assess the time cost in terms of the absolute difference between switch and nonswitch RT, it diminished by only 29% and was still very substantial (186 ms). But absolute cost may not best represent the effects of practice, because RT on nonswitch trials also diminished with practice (by 13%). Expressed as a proportion of this baseline, the decline in switch cost with an extra 1,000 trials of practice amounted to only 7%. We may therefore feel confident that the switch costs observed are not a phenomenon limited to a transient phase of acquisition of new task-sets or learning to switch between them. Moreover, the results of the transfer test establish that neither the improvement in performance with practice nor such decline in switch cost as there was could reasonably be attributed to the participant's coming to respond on the basis of specific learned S–R associations, rather than continuing to switch between the two classification tasks defined by the experimenter. (Of course, such a transition might well occur in time, with much more practice.)

The stimuli were designed to allow us to manipulate the association between task-irrelevant stimulus attributes and the now-irrelevant task—and thus to look for a task-cuing effect of the irrelevant attribute. Several features of the data indicate the presence of such an effect. First, performance was, overall, slower (by 173 ms) and less accurate in the crosstalk than in the no-crosstalk condition. Second, performance in the crosstalk condition was slower, and no more accurate, when the irrelevant attribute was associated with the other task but congruent (i.e., when it was mapped to the currently appropriate response via that task) than when it had no association with either task or response. Third and fourth, what is true of performance was also true of switch costs: Switching was harder in the crosstalk condition than in the no-crosstalk condition and harder with a congruent than with a neutral irrelevant character.

We have argued that to shift to a different task-set one must in some sense enable the particular S–R mapping appropriate for that task and suppress competing S–R links specified by the now-inappropriate task. The contrast between congruent and incongruent trials within the crosstalk condition revealed a modest crosstalk effect, more so for error rate than for RT, which suggests that the competing task-set was not entirely disabled, especially on the first trial after a task switch. However, for RT at least, the task-cuing effect was much stronger than this crosstalk effect. This clearly indicates that stimulus attributes activate not just the S–R association specified by a primed but irrelevant task, but the task-set itself: the whole package of S–R associations. This activation of a (primed) competing task-set interferes with the exercise of the required task-set, especially at the point where one is switching from that competing task. Moreover, in terms of the temporal cost of a switch, this effect substantially outweighs the benefit on those trials when the inappropriate S–R mapping happens to specify the production of the same response as the now-appropriate one.

This task-cuing effect indexes, we suggest, the same exogenous component of control that is seen in extreme form in capture errors in normal people and in the utilization behavior of frontal lobe patients. Stimuli tend to evoke the task-set associated with them. In the non-brain-damaged person this tendency can be overcome, but at some cost in time. Indeed, it is noteworthy that the main consequence of the task-cuing effect we observed in this experiment was to delay the response, not to generate errors.

Even in the no-crosstalk condition there was a substantial cost, in both time and errors, when the participant had to switch tasks. As we have already noted, a natural interpretation of the time cost is as the duration of a reconfiguration process. A physical metaphor would be the time taken by a signalperson to heave, from one position to another, the heavy lever that, in preelectronic railroad systems, is used to operate the switch (or "points" in British English) between railroad tracks so that the next train would run onto a different line. This switching-duration metaphor leads to a simple prediction. One cannot move the switch to Line B until after the previous train needing to go down Line A has passed the switch. If a train needing to go down Line B arrives before one has had time to move the lever, it will have to wait. That is the origin of the switch cost. However, if enough time elapses before the arrival of the next train, and the signalperson knows in advance that it is destined for Line B, there will be time to heave the lever across to the new position, and no delay should ensue. Indeed, by varying the time between successive trains needing to go down different lines and measuring the delay, one should be able to measure the duration of the lever-heaving process. This, in essence, was the rationale for the next three experiments.

Experiment 2

The hypothesis to be tested was that the temporal cost of switching between tasks observed in Experiment 1 reflects the duration of an endogenously controlled task-set reconfiguration process that can, if enough time is available, be performed in anticipation of the task-switch. The test was simply to increase the R–S interval to a value well beyond the longest possible duration of the reconfiguration process while maintaining the predictability of the task-switch. We

therefore varied the R–S interval between 150 ms (as in Experiment 1) and 1,200 ms. The hypothesis predicts that as the R–S interval is prolonged, the switch cost should decline, as more of the task-reconfiguring process is completed before the next stimulus is presented, and that with a long enough interval, it should reach zero. In a system with no variability, the function relating time cost to R–S interval should be linear with a slope of minus 1 until it reaches zero, and flat thereafter. Any variability will tend, however, to round the point of inflexion and make the slope shallower. (Compare the function relating stimulus onset asynchrony and RT in the "psychological refractory period" paradigm—e.g., Pashler & Johnston, 1989).

In our first attempt at this manipulation we elected to vary the interval randomly from trial to trial. The rationale was that, not being able to predict the interval, the little signal-person in the head would be encouraged to pull the reconfiguration lever as soon and as quickly as possible after every response preceding a task-switch.

Method

The tasks, stimuli, and equipment were the same as for Experiment 1. Five intervals were used: 150, 300, 450, 600, and 1,200 ms (150 msec is the shortest practicable R–S interval that can be used because of the need for the participant to release the key before a further keypress can be initiated and registered). We selected 1,200 ms as a value substantially greater than the maximum switch cost observed in Experiment 1.

On each of 2 days of testing, the participants completed 20 blocks of trials identical to the crosstalk condition of Experiment 1, apart from the manipulation of the R–S interval. Each block consisted of 48 trials, the first 12 of which were treated as warm-up trials and excluded from the data analysis. The 20 blocks contained one trial with every combination of the variables: task, trial type, R–S interval, response, irrelevant character type, response on the previous trial, and irrelevant character type on the previous trial.

The participants were allowed a break of about 10 min after the first 10 blocks. The preexperimental training was the same as in Experiment 1, except that in the two switching practice blocks, the R–S interval varied randomly from trial to trial. The participants were instructed to respond to each character pair as rapidly as possible while avoiding errors and then to try to prepare for the next trial. If this involved a task-switch, a participant was to ready himself or herself for the new task as quickly as possible following the previous response.

A new group of 8 paid participants (3 men and 5 women, between 19 and 38 years of age) was instructed to signal *consonant* and *even* decisions by pressing the c key of the computer keyboard with the index finger of the left hand and to signal *vowel* and *odd* decisions by pressing the m key with the index finger of the right hand.

Results

Switch cost. Mean RT and error proportions are shown in Table 3 (subject to the same exclusions as in Experiment 1). ANOVAs were performed with the factors of R–S interval (150, 300, 450, 600, and 1,200 ms), practice, task,

trial type, and irrelevant character type. As in Experiment 1, switching tasks reliably slowed RT, $F(1, 7) = 39.0, p < .001, MSE = 239,105$, the time cost being larger on the 1st day of testing (245 ms) than on the 2nd day (149 ms), $F(1, 7) = 13.4, p < .01, MSE = 41,144$. There was also a significant error cost, $F(1, 7) = 29.7, p < .01, MSE = 0.176$, but it was not reliably reduced between the 1st (3.7%) and 2nd days (2.5%), $F < 1$.

Effect of R–S interval. Switch cost is plotted as a function of R–S interval in Figure 4, along with the data from Experiments 3 and 4. To our initial surprise, there was no sign whatsoever of a reduction in the time cost of a switch as the R–S interval increased, $F < 1$. Analysis of simple effects demonstrated that the time costs were reliable at every R–S interval, $p < .05$ in each case. There was also no reliable reduction in error cost with increasing R–S intervals, $F < 1$. (In simple effects tests, the error cost fell short of a reliable level at an R–S interval of 300 ms, $F[1, 7] = 2.6$, and only approached significance at an interval of 450 ms, $F[1, 7] = 4.39, p = .07, MSE = 0.409$; the costs at all other R–S intervals were reliable, $p < .05$ in each case.) Finally, there was no sign of a three-way interaction among switching, R–S interval, and day to suggest that the expected relation between R–S interval and switch cost was developing with practice: $F(4, 28) = 1.7$ for the RTs, and $F(4, 28) = 1.9$ for the error rates. Mean latency was not reliably affected by R–S interval, $F(4, 28) = 1.3$. However, on the 1st day, RT was reduced by 21 ms as the R–S interval was prolonged from 150 to 450 ms; the same contrast on the 2nd day showed only a 2-ms reduction. This interaction between R–S interval and practice was marginally reliable, $F(4, 28) = 2.6, p = .06, MSE = 4,639$. Error rates were slightly lower at longer R–S intervals, $F(4, 28) = 2.6, p = .06, MSE = 0.054$, but there was no obvious interaction of this effect with practice, $F(4, 28) = 1.5$.

Effect of irrelevant character. As in Experiment 1, RTs were slower on both switch and nonswitch trials when the irrelevant character was associated with the now-inappropriate task, $F(2, 14) = 32.1, p < .001, MSE = 68,703$. Such crosstalk also increased overall error rates, $F(2, 14) = 14.9, p < .001, MSE = 0.276$. These effects did not interact with R–S interval, $F < 1$ for both performance measures.

The time cost of task-switches was significantly increased by irrelevant characters associated with the now-inappropriate task, $F(2, 14) = 5.6, p < .05, MSE = 12,261$. Thus, as in Experiment 1, the time cost was largest for the incongruent case (223 ms), somewhat less for the congruent case (203 ms), and smallest in the neutral case (165 ms). Only the differences between the neutral case and the other two were reliable ($p < .05$, in each case). The error cost was also affected by the irrelevant character, $F(2, 14) = 12.4, p < .001, MSE = 0.097$, and, as in Experiment 1, the pattern for the error cost was different from the pattern for the time cost: 0.8% for the congruent case, 8.4% for the incongruent case, and 2.2% for the neutral case. The error cost for the incongruent case was reliably larger than the error costs for the other two cases ($p < .01$). The interaction among R–S interval,

Table 3

Performance and Switch Costs by Response–Stimulus (R–S) Intervals and Tasks in Experiment 2

R–S interval	Letter		Digit		Average	
	RT	% error	RT	% error	RT	% error
150 ms						
Switch trials	905.7	6.8	933.0	7.3	919.3	7.0
Nonswitch trials	721.8	4.4	730.7	4.2	726.2	4.3
Switch cost	183.9	2.4	202.3	3.1	193.1	2.7
300 ms						
Switch trials	923.9	5.8	915.4	6.2	919.6	6.0
Nonswitch trials	762.6	4.7	706.0	4.2	734.3	4.4
Switch cost	161.3	1.1	209.4	2.0	185.3	1.5
450 ms						
Switch trials	910.9	5.6	910.2	8.3	910.5	6.9
Nonswitch trials	725.0	1.7	700.0	3.4	712.5	2.5
Switch cost	185.9	3.9	210.2	4.9	198.0	4.4
600 ms						
Switch trials	941.6	6.6	885.4	7.4	913.5	7.0
Nonswitch trials	724.1	3.2	692.3	3.1	708.2	3.1
Switch cost	217.5	3.4	193.1	4.3	205.3	3.8
1,200 ms						
Switch trials	925.8	6.1	923.0	5.6	924.4	5.8
Nonswitch trials	722.9	3.3	718.2	2.2	720.5	2.7
Switch cost	202.9	2.8	204.8	3.4	203.8	3.1

Note. RT = mean reaction time in milliseconds.

trial type, and irrelevant character was not significant for RTs, $F(8, 56) = 1.3$, or error rates, $F < 1$.

Discussion

This first attempt to measure the time course of any anticipatory reconfiguration of task-set evidently did not succeed. Neither the time nor the error measures of the cost of a predictable task-switch decreased as the R–S interval increased from 150 to 1,200 ms. Mean RT also showed little sign of the "aging foreperiod" effect often observed with a variable but bounded R–S interval (Bertelson, 1967; Bertelson & Tisseyre, 1968; Näätänen, 1970).

Could the random variation in R–S interval have discouraged anticipatory reconfiguration rather than—as we intended—encouraging it? If we return to our railroad metaphor, it is obvious that a sensible signalperson would not pull the switch lever if there was any possibility of a train's running through the switch while the switching operation was still in progress, because the train would derail! By analogy, our 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. To test this hypothesis we reran

Experiment 2 but the R–S interval held constant within a block so that the participant could predict when the next stimulus would arise.

Experiment 3

Method

On the 1st day of testing, the participant was first given extensive practice with each task alone, as before. He or she then completed two practice switching blocks with an R–S interval of 150 ms, followed by five sequences of four blocks, each sequence with a different R–S interval. On the 2nd day, the participant first completed two more practice blocks with an interval of 150 ms and then five more sequences of four blocks, again one with each R–S interval. The order of testing on the five R–S intervals was counterbalanced over 5 participants by means of a Latin square, with the order on the 2nd day the reverse of that on the 1st day. As before, the first 12 trials of each block were treated as warm-up trials, and data from them were discarded.

Each sequence of four blocks contained one trial with every combination of task, trial type, response, response on the previous trial, irrelevant character type, and irrelevant character type on the previous trial. A break was allowed between each sequence of four blocks and the next, and the participant was informed about the length of the next R–S interval.
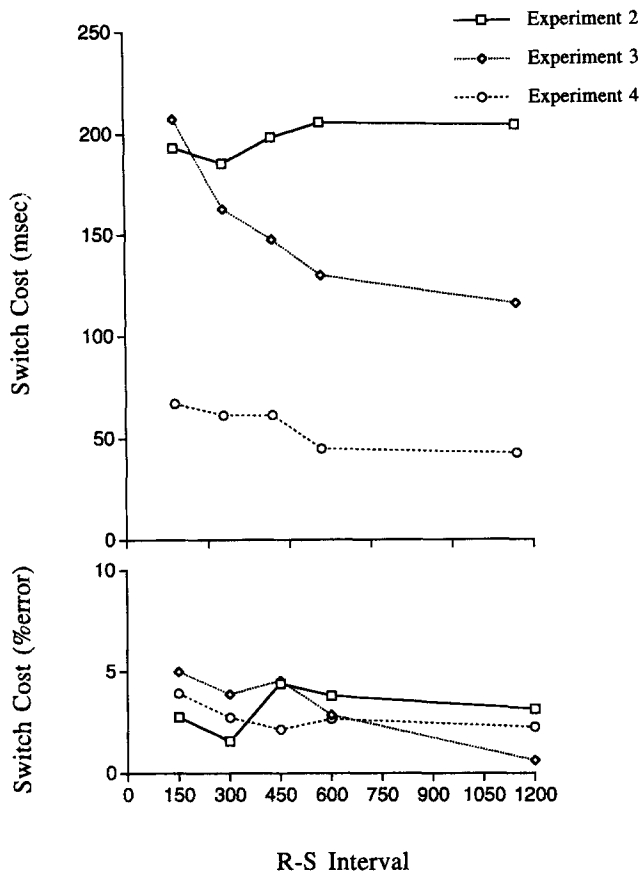
Figure 4. The cost in mean reaction time and error rate of switching task (that is, the difference between performance on switch and nonswitch trials), as a function of response–stimulus (R–S) interval, in Experiment 2 (varied R–S interval, crosstalk condition), Experiment 3 (constant R–S interval, crosstalk condition), and Experiment 4 (constant R–S interval, no-crosstalk condition), averaged over task and day.

There were 10 new participants, 4 men and 6 women, between 18 and 30 years of age. Other details were as in Experiment 2.

Results

Figure 4 shows that, as the constant R–S interval used in Experiment 3 was increased from 150 ms, the time cost of a switch declined, leveling out between the R–S intervals of 600 and 1,200 ms. The effect of R–S interval on the time cost was highly reliable, $F(4, 36) = 6.4, p < .001, MSE = 11,704$. The cost at the 150-ms R–S interval was reliably larger than that at the 300-ms and 450-ms intervals ($p < .05$), and at the 600-ms and 1,200-ms intervals ($p < .01$). Hence it is clear that the time costs incurred by predictable task-switches can be reduced by prolonging the interval available for preparation between the participant's last response and the next stimulus. However, it is equally clear that switch cost did not decrease to zero: A large asymptotic switch cost remained when the R–S interval was extended past 600 ms to 1,200 ms. Analysis of simple effects showed

that the time cost at every R–S interval was reliable, $p < .05$ in all cases.

Prolonging the R–S interval also steadily reduced the error cost, $F(4, 36) = 2.7, p < .05, MSE = 0.150$. However, in contrast to the time cost, the error cost continued to diminish between the R–S intervals of 600 and 1,200 ms. A Newman-Keuls test showed that only the differences in error cost between the 150-ms (5.0%) and the 1,200-ms (0.6%) RS intervals reached significance, $p < .05$. There was also a marked increase in errors on nonswitch trials as R–S interval was increased to 600 ms and then 1,200 ms—resulting in a reliable simple effect of R–S interval for nonswitch trials, $F(4, 36) = 3.9, p < .01, MSE = 0.075$—so that the decrease in the error cost with increasing R–S interval was not caused only by an improvement in accuracy on the switch trials.

Mean RT also decreased significantly as the R–S interval increased from 150 to 600 ms (see Table 4), $F(4, 36) = 5.4, p < .005, MSE = 34,334$. Newman-Keuls tests demonstrated that RT at the 150-ms R–S interval was reliably longer than that at every other R–S interval ($p < .05$ in each case). There was no consistent decrease in error rates over the entire range of R–S intervals, although the main effect was almost reliable, $F(1, 9) = 2.1, p = .09, MSE = 0.073$. Switching tasks increased RT more on the 1st day than the 2nd day (186 vs. 119 ms), $F(1, 9) = 18.0, p < .005, MSE = 18,935$. The corresponding change in the error cost with practice, from 4.1% on the 1st day to 2.7% on the 2nd day, was not reliable, $F(1, 9) = 2.2, p > .1, MSE = 0.165$.

As before, a marked effect of the irrelevant character type (congruent, incongruent, or neutral) was observed: This variable influenced mean RT, $F(2, 18) = 59.0, p < .001, MSE = 25,246$; error rate, $F(2, 18) = 7.4, p < .005, MSE = 0.280$; the time cost of a task-switch, $F(2, 18) = 18.4, p < .001, MSE = 4,316$; and the error cost of a switch, $F(2, 18) = 4.8, p < .05, MSE = 0.119$. Both overall performance and cost incurred by task-switches were worse when the irrelevant character was non-neutral. The time cost was 156 ms on congruent trials, reliably smaller than the 179-ms cost on incongruent trials ($p < .05$), whereas both of these costs were larger than the 123-ms cost on neutral trials ($p < .01$).

There was no significant two-way interaction between R–S interval and irrelevant character type for mean RT, $F < 1$. However, there was a marginally reliable interaction for error rate, $F(8, 72) = 1.8, p = .07, MSE = 0.090$, the simple effect of R–S interval being reliable on incongruent trials only, $F(4, 36) = 2.8, p < .05, MSE = 0.091$. There was also no reliable interaction between R–S interval, irrelevant character type, and trial type for either the RTs or error rates, $F < 1$ in both cases.

Discussion

Keeping the R–S interval constant appears to have allowed and encouraged the participant to do something useful in anticipation of a task-switch when enough time was allowed to do so: A progressive saving of the switch cost

Table 4

Performance and Switch Costs by Response–Stimulus (R–S) Intervals and Tasks in Experiment 3

	Letter		Digit		Average	
R–S interval	RT	% error	RT	% error	RT	% error
150 ms						
Switch trials	835.6	7.8	879.3	7.5	857.4	7.6
Nonswitch trials	657.0	1.8	643.5	3.5	650.2	2.6
Switch cost	178.6	6.0	235.8	4.0	207.2	5.0
300 ms						
Switch trials	763.4	6.5	799.4	6.7	781.4	6.6
Nonswitch trials	625.5	2.5	612.3	2.9	618.9	2.7
Switch cost	137.9	4.0	187.1	3.8	162.5	3.9
450 ms						
Switch trials	743.9	7.4	785.5	9.2	764.7	8.3
Nonswitch trials	621.4	3.4	613.3	4.2	617.3	3.8
Switch cost	122.5	4.0	172.2	5.0	147.3	4.5
600 ms						
Switch trials	733.1	5.9	774.2	7.5	753.6	6.7
Nonswitch trials	637.6	3.0	610.1	4.7	623.8	3.8
Switch cost	95.5	2.9	164.1	2.8	129.8	2.8
1,200 ms						
Switch trials	724.7	5.5	769.0	5.6	746.8	5.5
Nonswitch trials	646.2	4.7	617.4	5.2	631.8	4.9
Switch cost	78.5	0.8	151.6	0.4	115.0	0.6

Note. RT = mean reaction time in milliseconds.

was observed as the R–S interval was increased from 150 ms to 600 ms. This much is compatible with the idea of a reconfiguration process that can be carried out in advance of the stimulus and may take as much as half a second to accomplish. However, only about one third of the switch cost was eliminated by preparation in this way. A substantial residual time cost remained, which resisted further reduction as the R–S interval was extended to more than 1 s (though the error cost was almost eliminated by a sufficient interval for preparation, albeit as much by an increase in errors on nonswitch trials as by a reduction on switch trials).

These data are in broad agreement with those from Experiment 5 of Allport et al. (1994) in suggesting two components of switch cost: one that is eliminated by an adequate R–S interval and a residual component that is not. Using Stroop stimuli, and a constant R–S interval, in two-trial sequences, with the participant warned before each sequence what the two tasks would be, Allport et al. found a (nonsignificant) reduction in the time cost of a switch as the R–S interval was increased from 20 to 550 ms but no further decrease as the interval was increased to 1,100 ms. Before we address possible explanations, we present data from two further experiments that, among *other things*, establish the robustness of this residual component of the switch cost.

Experiment 4

On two thirds of the trials in Experiment 3, the irrelevant character in the display was associated with the now-inappropriate task, as in the crosstalk condition of Experiment 1. Might the residual component of the switch cost be associated specifically with the inhibitory component of task-set—that is, with the need to suppress crosstalk? To find out, we simply reran Experiment 3 but with irrelevant characters drawn entirely from the neutral set, as in the no-crosstalk condition of Experiment 1.

Method

Experiments 3 and 4 differed in only two respects. First, the irrelevant character was always drawn randomly from the neutral set (as in the no-crosstalk condition of Experiment 1). Second, each block consisted not of 48 but of 40 trials, of which only the first 4, rather than the first 12, were discarded as warm-up trials. Analysis of the earlier experiments had shown that there was no obvious change in performance through the block and that a long lead-in was unnecessary. The 10 new participants (7 men and 3 women) were between 19 and 28 years of age.

Results

Mean RTs and error rates are presented in Table 5, and the switch costs are plotted as a function of R–S interval in Figure 4. The time cost of a task-switch, $F(1, 9) = 59.2$, $p < .001$, $MSE = 5,172$, and the error cost, $F(1, 9) = 20.8$, $p < .005$, $MSE = 0.096$, were both highly reliable. Comparing these costs with those observed in Experiment 3, we found that the time cost was substantially smaller, $t(18) = 2.5$, $p < .05$, although the error cost was not, $t(18) = 0.5$. This replicates between subjects the contrasts between crosstalk and no-crosstalk conditions observed within subjects in Experiment 1. However, as in Experiment 3, the time cost decreased as the R–S interval increased, $F(4, 36) = 5.9$, $p < .005$, $MSE = 412$, from 67 ms at an interval of 150 ms to 45 ms at an interval of 600 ms and to 43 ms at an interval of 1,200 ms. Thus, as in Experiment 3, an increased R–S interval allowed the participants to reduce the time cost of a switch by about one third, but not to eliminate it altogether. The simple effect of the switch was still reliable at an R–S interval of 1,200 ms, $F(1, 9) = 47.5$, $p < .001$, $MSE = 760$. The error cost of a switch also declined with an increasing R–S interval, but unreliably so, $F < 1$; it was still 2.2% at an interval of 1,200 ms.

As in Experiment 3, mean RTs on both switch and nonswitch trials were smaller at longer R–S intervals than at shorter ones (see Table 5), $F(4, 36) = 21.3$, $p < .001$, $MSE = 1,112$. Newman-Keuls tests showed that mean RT at R–S intervals ≥ 300 ms was reliably faster than that at the 150-ms interval, $p < .01$. RTs were also significantly smaller at the 600-ms R–S interval than at intervals ≥ 150 ms, $p < .05$. Practice reduced this main effect of R–S interval, $F(4, 36) = 2.5$, $p = .05$, $MSE = 3,118$, but analysis of the simple effects showed that R–S interval was still a significant source of variance on the 2nd day of testing, $F(4, 36) = 7.8$, $p < .001$, $MSE = 843$. There was a reliable main effect of R–S interval for error rates as well, $F(4, 36) = 2.7$, $p < .05$, $MSE = 0.033$, although this arose from an increase between the 600-ms and 1,200-ms R–S intervals.

The time cost of a switch was reduced with practice from 70 ms on the 1st day to 41 ms on the 2nd day, $F(1, 9) = 33.2$, $p < .001$, $MSE = 634$, and the error cost fell from 3.6% to 1.9%, $F(1, 9) = 8.9$, $p < .05$, $MSE = 0.024$. However, there was no reliable three-way interaction among switching, R–S interval, and practice for either performance measure.

Discussion

In general, Experiment 4 replicated the findings of Experiment 3, with proportionately smaller switch costs attrib-

Table 5
Performance and Switch Costs by Response–Stimulus (R–S) Intervals and Tasks in Experiment 4

R–S interval	Letter		Digit		Average	
	RT	% error	RT	% error	RT	% error
150 ms						
Switch trials	570.0	8.5	563.3	6.8	566.6	7.6
Nonswitch trials	501.5	3.1	497.8	4.4	499.6	3.7
Switch cost	68.5	5.4	65.5	2.4	67.0	3.9
300 ms						
Switch trials	540.7	7.2	527.9	7.3	534.3	7.2
Nonswitch trials	472.4	2.9	473.5	6.2	472.9	4.5
Switch cost	68.3	4.3	54.4	1.1	61.3	2.7
450 ms						
Switch trials	537.4	7.2	524.8	6.7	531.1	6.9
Nonswitch trials	465.0	3.1	475.0	6.4	470.0	4.7
Switch cost	72.4	4.1	49.8	0.3	61.1	2.2
600 ms						
Switch trials	510.4	8.1	506.8	8.2	508.6	8.1
Nonswitch trials	460.3	5.5	467.3	5.5	463.8	5.5
Switch cost	50.1	2.6	39.5	2.7	44.8	2.6
1,200 ms						
Switch trials	525.6	6.6	518.3	5.8	521.9	6.2
Nonswitch trials	476.4	3.9	482.5	4.0	479.4	3.9
Switch cost	49.2	2.7	35.8	1.8	42.5	2.2

Note. RT = mean reaction time in milliseconds.

utable to the lack of any possibility for crosstalk between tasks. The time cost was once again reduced by about one third as the R–S interval was prolonged up to about 600 ms, but no further decrease occurred and a stubborn residual cost remained even when the R–S interval was as long as 1,200 ms. The error cost also declined (albeit unreliably) as the R–S interval increased, but in this experiment it remained appreciable at the longest R–S interval. Hence the resistance of the temporal cost of a switch to allowing the participant ample time for anticipatory reconfiguration does not appear to be limited to the case in which the stimulus frequently (Experiment 3) or always (Allport et al., Experiment 5) contains attributes associated with the now-irrelevant task. The residual component of the switch cost does not arise only from the need to suppress crosstalk.

The experiment we report next disposes of a further hypothesis that, though paradoxical, must nevertheless be taken seriously. This hypothesis proposes that at long R–S intervals we were, in a sense, allowing the subject too much time to reconfigure: Temporal uncertainty was interfering with optimal preparation.

Experiment 5

It has long been argued, from foreperiod effects in RT experiments, that achieving a state of optimal preparation for the reaction signal is effortful and that maintaining that state is aversive. In simple RT experiments with a varied foreperiod and rectangular distribution of delays, RT tends to decline toward the end of the range of intervals, which suggests that, as the probability of the signal's occurring in the next small unit of time becomes higher, the participant becomes readier to respond, rather than getting maximally ready at the beginning of the interval and maintaining that level of preparation (Näätänen, 1970). In experiments with constant foreperiods, it has been found that as the foreperiod exceeds 0.5 s or so, RT shows an upward trend, for both simple RT (Klemmer, 1956; see Näätänen & Merisalo, 1975, for a review) and choice RT (Bertelson, 1967; Bertelson & Boons, 1960; Bertelson & Tisseyre, 1968, 1969). Rather than maintaining an optimum level of preparation throughout the foreperiod, participants appear to target their preparation around the reaction signal's expected time of occurrence. However, because of the Weber fraction for time estimation, with longer intervals, prediction of the time of occurrence of the imperative signal becomes less precise. This makes the transient "peak" of preparation harder to target appropriately. The best balance between the benefits of adequate time to prepare, the aversion to maintaining preparation, and the loss of temporal uncertainty with an increase in foreperiod is, it seems, generally found at around half a second (Gottsdanker, 1975).

This concept of temporally targeted preparation can be readily extended to the preparation involved in reconfiguring for a change of task. The hypothesis is that, in Experiments 3 and 4, the switch cost declined up to R–S intervals *of 600 ms because the reconfiguration process took several hundred milliseconds, but it could be relatively precisely*

targeted for the arrival of the next stimulus. However, as the interval was extended to 1,200 ms, the deterioration in temporal certainty impaired this targeting. Participants tended either to reconfigure too soon, and were unable to maintain an appropriate task-set indefinitely, or to reconfigure too late. It is concordant with this hypothesis that in Experiment 2, with a highly variable R–S interval, participants apparently declined to prepare at all, and that in Experiments 3 and 4, RT on nonswitch trials was optimal at intermediate R–S intervals.

To test the hypothesis, we simply reduced temporal uncertainty at a long R–S interval by adding an otherwise uninformative warning signal. In Experiment 5, the R–S interval was a constant 1,000 ms throughout, and there were two (blocked) conditions. In the warning-signal condition, an uninformative warning signal preceded each stimulus by 500 ms, whereas in the no-warning-signal condition, there was no such warning signal.

Method

The method was essentially the same as that of Experiment 4, with the following differences. On the 1st day, after preexperimental training, half of the participants completed two practice blocks and then two four-block sequences in the no-warning-signal condition, followed immediately by two four-block sequences in the warning-signal condition. The other participants first completed two practice blocks and two sequences in the warning-signal condition, immediately followed by two sequences in the no-warning-signal condition. On the 2nd day of testing, these orders were reversed. The preexperimental training was under the same warning-signal condition as the initial experimental blocks. There were 8 participants (4 men and 4 women) between 19 and 30 years of age.

The warning signal consisted of a box (subtending 2.2° horizontally and 1.1° vertically) surrounding the locus of the next stimulus. The signal was presented 500 ms prior to stimulus onset and remained visible concurrently with the stimulus. (It may be objected that this warning signal was informative in the sense that it enabled more precise fixation on the locus of the next stimulus. Fortunately, given the outcome, we need not counter this objection.)

Results

Mean RT and error rates were examined with ANOVAs having the factors of practice, condition (no warning signal vs. warning signal), task, and trial type. As may be seen in Table 6, switching tasks in this experiment incurred a time cost of 61 ms, $F(1, 7) = 11.7, p < .05, MSE = 10,216$, and an error cost of 2.3%, $F(1, 7) = 10.3, p < .05, MSE = 0.046$. The time cost was reliably reduced by practice, $F(1, 7) = 7.4, p < .05, MSE = 1,803$, although the simple effect of switching on the 2nd day of the experiment (41 ms) remained highly reliable, $F(1, 7) = 14.7, p < .01, MSE = 1,795$.

The provision of a warning signal reduced mean RT by 13 ms, $F(1, 7) = 6.6, p < .05, MSE = 819$, but at the cost of a 1.4% increase in error rates, $F(1, 7) = 6.7, p < .05, MSE = 0.024$. On the other hand, presenting a warning

Table 6
Performance and Switch Costs by Warning-Signal Conditions and Tasks in Experiment 5

Condition	Letter		Digit		Average	
	RT	% error	RT	% error	RT	% error
No warning signal						
Switch trials	539.5	5.0	558.7	5.5	549.1	5.2
Nonswitch trials	481.7	3.0	489.2	3.7	485.4	3.3
Switch cost	57.8	2.0	69.5	1.8	63.6	1.9
Warning signal						
Switch trials	529.0	6.1	537.9	8.0	533.4	7.0
Nonswitch trials	472.3	4.7	477.8	4.0	475.0	4.3
Switch cost	56.7	1.4	60.1	4.0	58.4	2.7

Note. RT = mean reaction time in milliseconds.

signal prior to the stimulus onset did not significantly reduce either the RT or error costs of a task-switch. The extra time needed to respond on switch trials was 64 ms without the warning signal and 58 ms with it, $F < 1$, whereas the provision of a warning signal increased the error cost by 0.8%, $F < 1$.

There was no reliable three-way interaction among warning condition, trial type, and task for the RTs, $F(1, 7) = 1.1$. Although the warning signal produced a small decrease in the error cost of switching to the letter task, it produced an increase from 1.8% to 4.0% in the error cost for the digit task, $F(1, 7) = 9.1$, $p < .05$, $MSE = 0.005$. Furthermore, practice interacted with the factors of task and trial type so that both tasks eventually showed marginal increases in the error cost with the warning signal on the 2nd day, $F(1, 7) = 4.6$, $p = .07$, $MSE = 0.022$.

Discussion

The small but reliable effect of the warning signal on mean RT and error rate indicates that participants registered and made use of the warning signal. The pattern of effects is consistent with other evidence that warning signals affect performance by inducing participants to respond sooner but on the basis of poorer quality information (Posner, Klein, Summers, & Buggie, 1973). The effect of the warning signal may seem small, but it is comparable to previous observations: Visual warning signals (Bertelson & Tisseyre, 1969) generally produce smaller facilitatory effects than auditory (Bertelson, 1967; Bertelson & Tisseyre, 1968) and tactile ones (Frith & Done, 1986). That this warning signal, while influencing overall performance, did not reliably modulate the switch cost suggests that the switch cost cannot be attributed to the temporal uncertainty associated with long foreperiods.

It is worth pointing out that, among the experiments reported in this article, this one should have provided optimal conditions for anticipatory configuration. The R–S interval remained constant throughout; the participants received much practice with it; it was surely long enough for

any anticipatory reconfiguration to take place; and an extra warning signal signaled its imminent end. Nevertheless a reliable cost of switching tasks, comparable to that observed at the longer R–S intervals in Experiment 4, was obtained. Participants' inability to use a preparatory interval to get themselves fully into the state they would be in had they just performed the same task appears to be a robust phenomenon.

Three Hypotheses for the Residual Switch Cost

Experiments 3, 4, and 5 demonstrated, as did Allport et al.'s (1994) Experiment 5, that there is a substantial component of the switch cost that cannot be eliminated by allowing the participant time to prepare for the task-switch. We now consider three hypotheses concerning this residual component of switch cost and their predictions for what should be observed when we extend the run of trials between task-switches beyond two, as we did in the final experiment we report.

The Task Set Inertia (TSI) Hypothesis

Allport et al.'s (1994) proposal, already reviewed in our introduction, was that the cost of switching arises because of interference exerted proactively by the residue of the previous task-set. In their words,

> The time costs of shifting task set that we observe in these experiments cannot be understood as the reflection of a discrete stage that must be completed before execution of the next S–R task can begin. Rather, like many other RT interference or conflict phenomena, they appear to represent the additional time needed for the system to settle to a unique response decision (or response retrieval) after the next imperative stimulus has arrived. (Allport et al., 1994, p. 436)

A crucial feature of Allport et al.'s account is that the TSI responsible for this additional time has a persistence on the order of "a minute or so" (p. 436). This is essential for explaining the gradual diminution of switch cost observed,

in Allport et al.'s Experiment 4, over successive blocks following a change between the tasks afforded by the two stimulus ensembles. Hence, if we increase the length of the runs in our alternating runs paradigm, not only the first trial on Task B following a switch from Task A, but a number of trials thereafter, should still show the effects of TSI from Task-set A, although we might expect the effect to diminish slowly through the run as the TSI is dissipated by the performance of Task B.

The Micro-Practice Hypothesis

Though one typically thinks of practice as being a long-term process, there are also dynamic processes of optimization that happen on a time scale of individual trials. For example, over a series of trials, choice RT reduces steadily until the participant makes an error, lengthens abruptly on the posterror trial, and then improves steadily again until another error is made (Rabbit & Vyas, 1970). This "scalloped" pattern of RT relative to error trials implies a dynamic tracking process for staying close to the optimum speed–accuracy trade-off. It also seems likely that task-specific categorization and decision criteria are also adjusted dynamically from trial to trial (e.g., Thomas, 1973; Thomas & Legge, 1970). The micro-practice hypothesis explains the costs of task-switching by assuming that the participant embarks on each new run of trials at even a familiar task with these various criteria set to values that are suboptimal: It takes a few trials to reoptimize performance on the basis of trial-by-trial feedback. Concordant with this idea is the observation that repeated exercise of the same S–R mapping (Duncan, 1975; Rabbit & Vyas, 1973) facilitates performance over and above the repeated identification of a particular stimulus (Bertelson, 1965) or the repeated production of a particular response (Rabbit, 1968). Like the TSI hypothesis, this micro-practice hypothesis appears to predict that the costs of switching tasks should take more than a single trial with the new task to dissipate.

The Stimulus-Cued Completion Hypothesis

This hypothesis proposes that an endogenous act of control deployed before the onset of the stimulus can achieve only part of the process of task-set reconfiguration. Completion of the reconfiguration is triggered only by, and must wait upon, the presentation of a task-associated stimulus. Moment-by-moment reorganizations of our processing resources to accomplish different tasks are, it is assumed, achieved by the combined and (often) sequential operation of (anticipatory) endogenous and (stimulus-triggered) exogenous control mechanisms. The former explains the diminution in switch costs with increasing R–S intervals, the latter the irreducible residuum.

The assumption that a task-relevant stimulus is needed to complete task-set reconfiguration, although post hoc, is compatible with several observations. First, utilization behavior in frontal lobe patients and capture errors in nonpatients indicate that stimulus attributes associated with Task

A tend to evoke the performance of Task A. Second, in Experiments 1 through 3, we saw that both performance of Task B and switching to Task B are made harder by the presence in the stimulus of attributes associated with (recently exercised) Task A. We propose that the same association between stimulus attributes and task-set that results in a tendency to perform a task when we do not intend to perform it is also effective in completing the process of reconfiguring for that task when we do want to perform it.

This hypothesis makes a strong prediction different from that made by either the task set inertia or micro-practice theories. After the stimulus-cued completion of task-set reconfiguration has been triggered once by an effective task-relevant stimulus, it should not need to be done again while the task remains unchanged. Hence the cost of switching between tasks should be detectable only on the first trial of a run in the new task. Experiment 6 was a test of this prediction.

Experiment 6

In Experiment 6, we used a constant R–S interval of 450 ms—long enough that any switch cost observed should reflect mainly the persistent component of switch cost. Instead of switching between the letter and digit tasks after runs of only two trials, the participant switched, still completely predictably, after a run of four trials in each task. Again, cues were provided to help the participant keep track of which task was required and the current position in a run of four.

Method

The background display now consisted of a circle divided into eight equal segments like slices of a pizza. The circle was also segregated into two halves by a thickened black line along the horizontal diameter (for half of the participants) or the vertical diameter (for the remaining half). Character pairs appeared, 3.5 cm from the center of the circle, centered in successive segments, in a clockwise sequence around the clock-face, 450 ms after the previous response (except after an error). Each participant was told that when the character pair appeared in a specified half of the clock-face (e.g., left or upper), he or she was to perform the letter task, and when it appeared in the other half, the digit task. Eight participants were tested (4 men and 4 women, between 18 and 38 years of age), 2 for each assignment of halves of the display to tasks. Hence, as successive character pairs were displayed clockwise in successive segments, every participant switched task predictably on every fourth trial (but in different segments of the display so that we could control for the direction of eye movements).

Testing was conducted over 2 days and consisted of sequences of two blocks each. Each sequence included one trial with each of the combinations of task, position in a run, response, and the three types of irrelevant character. New sequences of trials were generated for every participant and day of testing. Each block consisted of 32 trials, the first two runs of four being treated as warm-up trials. On each day of testing, the participant completed two practice blocks and then 10 sequences of two blocks. A rest of 2 or 3 min was allowed after every two sequences. The response

assignments were the same as those used in Experiment 1, and the usual pretraining was given.

Results

Mean RTs and error rates are shown in Figure 5 as a function of position in a run of four trials on a given task. The result is unambiguous: Performance improves markedly from the first trial (i.e., the switch trial) to the first postswitch trial but does not improve thereafter.

The data were submitted to ANOVAs with the factors of practice, task, position in a run (Trial 1, 2, 3, or 4), and type of irrelevant character. The effect of position in a run was highly significant for RTs, $F(3, 21) = 11.6$, $p < .001$, $MSE = 100,390$. Newman-Keuls comparisons showed that the mean RT on Trial 1 of a run (i.e., with switch trials) was reliably longer than that on Trials 2, 3, or 4, $p < .01$. No other difference was significant. Inasmuch as there was any trend in RT over Trials 2 through 4, it was a 10-ms increase

rather than a decrease. However, an analysis excluding Trial 1 confirmed that this was not significant, $F < 1$.

The effect of position in a run was also highly reliable for the error rates, $F(3, 21) = 5.1$, $p < .01$, $MSE = 0.142$. Newman-Keuls tests showed that the error rate on Trial 1 was significantly higher than on Trials 2 and 3, $p < .05$; no other differences were reliable. As with the RTs, there was no sign of an improvement in accuracy as the participants continued with the same task from Trials 2 through 4 of a run; if anything, there was a trend in the other direction, though again not a reliable one, $F(2, 14) = 1.3$. As expected, practice somewhat reduced both the time cost associated with a task-switch, $F(3, 21) = 6.5$, $p < .005$, $MSE = 3,528$, and the error cost, the latter not reliably, $F < 1$.

RTs on switch and nonswitch trials were significantly increased when the irrelevant character was associated with the other task, $F(2, 14) = 26.4$, $p < .001$, $MSE = 14,152$, and RTs on both congruent trials (672 ms) and incongruent trials (667 ms) were reliably longer than those on neutral trials (576 ms), $p < .05$. The absence of an overall advantage for response times on congruent trials over incongruent trials was due to a marginally reliable two-way interaction between irrelevant character and position in a run, $F(6, 42) = 2.1$, $p = .07$, whereby RT was faster on congruent trials than on incongruent trials on Trials 1 and 4, but not on Trials 2 and 3. Accuracy on congruent and incongruent trials was also slightly worse than on neutral ones (6.1% and 5.7% vs. 5.0%), though unreliably so, $F(2, 14) = 2.2$. And, as with the RTs, the two-way interaction between irrelevant character and position in the run approached significance, $F(6, 42) = 2.0$, $p = .08$, $MSE = 0.128$, so that accuracy tended to be improved on congruent trials compared with incongruent trials on the first and last trials of a run, but not on the intervening trials. Finally, practice interacted significantly with the effect of the irrelevant character on RTs, $F(2, 14) = 6.8$, $p < .01$, $MSE = 2,046$, but not on error rates, $F < 1$.

Discussion

The outcome of Experiment 6 was clear-cut. Performance improved radically between the first and second trials of a run and then showed no sign of further improvement. In other words, the residual component of the cost of a task-switch did not dissipate gradually over a run of trials on the new task, but was limited to the first trial. This outcome appears incompatible with the TSI account of Allport et al. (1994), because it is essential to the explanatory adequacy of TSI in their account that it persist over many intervening trials. The outcome is also incompatible with the hypothesis that attributes switch cost to the need to reoptimize performance on every new run of trials in a task by micro-practice, unless it is assumed that this reoptimization can be reliably accomplished in just one trial. This seems implausible: If memory for the previous optimal settings was accurate enough to allow reoptimization in a single trial, why should it not be accurate enough for no trials to be needed? The outcome is, on the other hand, as predicted by the hypoth-
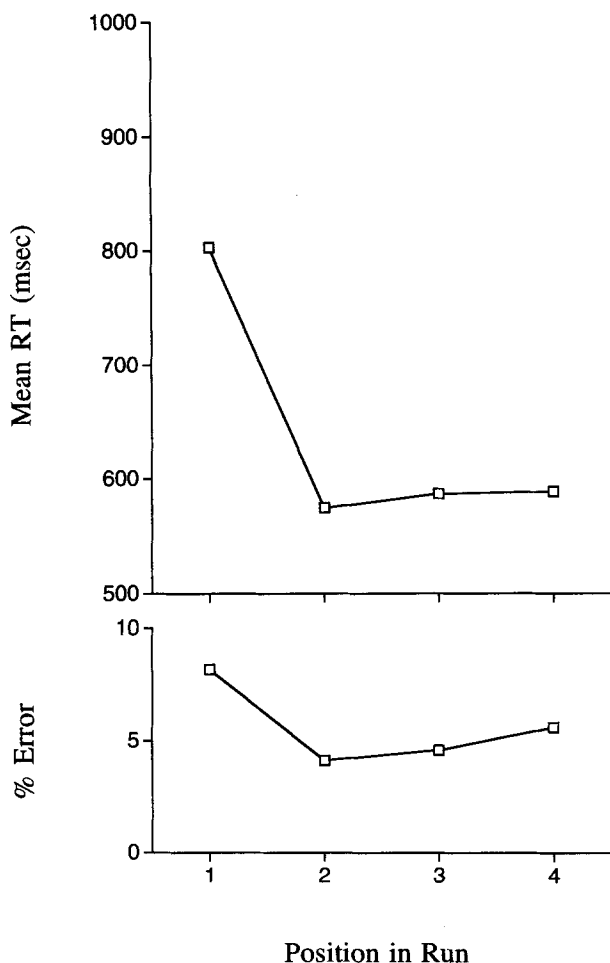

Figure 5. Mean reaction time (RT) and error rate in Experiment 6, as a function of position in a run of four trials on a task, averaged over task and day.

esis that, although task-set reconfiguration can be initiated endogenously, the exogenous trigger of a stimulus attribute associated with a task is needed to complete the process of reconfiguring to perform that task. Before we proceed to some general conclusions, we discuss two further issues on which analyses covering the whole set of experiments may shed some light.

Interaction of Response Repetition Priming and Task-Switching

It is commonly observed that, within a series of trials on the same task, RT and error rate are reduced when the response happens to be the same as the response produced on the previous trial (e.g., Bertelson, 1965; Pashler & Baylis, 1991; Rabbit, 1968). However, analysis of response repetitions in the present experiments yielded the unexpected finding that this advantage of response repetition did not apply when there was an intervening task-switch. Table 7 summarizes the effects of response repetition in all six experiments. It shows that in every case, mean RT on nonswitch trials was significantly reduced when the response was the same as on the previous trial. Error rate was also smaller, though not always reliably so. Thus, within a run, we observed the usual benefit of response repetition. In contrast, after a task-switch, there was no facilitation of RT when the same response was required as on the previous trial—indeed, in one experiment significant inhibition was observed. Moreover, there was also, in every case except one, a significant increase in the probability of an error when a response was repeated after a task-switch.

Table 7

Benefit of Response Repetition as Shown by Mean Performance for Alternated Response on Successive Trials Minus Mean Performance for Repeated Response

Experiment	Nonswitch trials	Switch trials	Interaction test	
Experiment 1				
RT	80**	−56*	$F(1, 7)$	= 35.2***
% error	0.2	−6.0**	$F(1, 7)$	= 8.8*
Experiment 2				
RT	57**	1	$F(1, 7)$	= 14.9**
% error	1.4	−4.7**	$F(1, 7)$	= 16.5**
Experiment 3				
RT	37**	−16	$F(1, 9)$	= 16.4**
% error	1.6	3.3**	$F(1, 9)$	= 20.8**
Experiment 4				
RT	59***	2	$F(1, 9)$	= 65.3***
% error	2.7*	−2.1**	$F(1, 9)$	= 35.8***
Experiment 5				
RT	37**	−1	$F(1, 7)$	= 15.2**
% error	1.6	−2.4*	$F(1, 7)$	= 15.6**
Experiment 6[a]				
RT	43*	−20	$F(3, 21)$	= 6.3**
% error	1.3	−1	$F(3, 21)$	= 1.7

Note. RT = mean reaction time in milliseconds.
[a] Nonswitch performance for Experiment 6 is averaged over Trials 2–4 of a run—see text.
*$p < .05$. **$p < .01$. ***$p < .001$.

Is this effect modulated by R–S interval? It will be recalled that switch cost was reduced with an increasing R–S interval in Experiments 3 and 4. In Experiment 3, there was, however, no reliable three-way interaction among R–S interval, switching, and response repetition, $F < 1$ for both RTs and errors. The RT data of Experiment 4 did yield a reliable three-way interaction, $F(4, 36) = 3.9$, $p < .05$, $MSE = 2,500$, but this was attributable to a decrease in the facilitation found for repeated responses on nonswitch trials after the longer R–S intervals. The interaction was not significant for the error rates, $F < 1$. Figure 6 shows a composite plot of the within-task and cross-task response repetition effects of all six experiments as a function of R–S interval. For response repetitions within a task, there is some suggestion that the RT benefit of a response repetition declines as the interval between successive responses increases. For response repetitions across a task-switch, there is some suggestion that when R–S interval is as short as 150 ms, and constant (i.e., in Experiments 1, 3, and 4), RTs to repeated responses are prolonged and the probability of an error is increased.

Why should a task-switch not only abolish the usual advantage enjoyed by a repeated response but also reverse it—at least in terms of error rate, and perhaps RT, at short lags? We offer three hypotheses:

1. The strengths of associations between the categorical attributes of a stimulus (i.e., consonant, vowel, even, odd) and motor responses (e.g., left or right index finger flexions) are continuously modified by experience. We may suppose that every response to a stimulus results in an increment in the associative strength of the link between the response and the stimulus category responsible and a corresponding decrement in the associative strength of the link between the other stimulus attributes and the response. These increments then decay. Production of a response triggered by the same attribute will be temporarily facilitated, whereas production of a response triggered by a different attribute should exhibit interference. Of course, this account appears to predict symmetrical facilitation (within task) and interference (across tasks), which is not quite what we observe: The facilitation is expressed mainly in RT, whereas the interference is expressed mainly in error rate. Another prediction made by this account is that, as long-term learning proceeds and the increments in associative weight attenuate as they approach asymptote, so the size of both facilitation and interference effects should diminish. In several of the experiments reported here (Experiments 2, 3, 4, and 6), both effects were reduced on the 2nd day compared to the 1st, and the three-way interaction among switching, response (repeated or unrepeated), and practice reached or approached significance in two of these: for errors rates in Experiment 3, $F(1, 9) = 5.7, p < .05, MSE = 0.041$, and for RT in Experiment 4, $F(1, 9) = 4.3, p = .07, MSE = 1,830$.

2. The effect reflects the operation of a control mechanism by which shifts of task-set are accompanied by the transient suppression of all active responses. That is, when a task-switch is called for, all ongoing response activity is inhibited as belonging to the previous task, and this inhibition takes some time to wear off. If a particular finger movement subject to this inhibition happens to be called
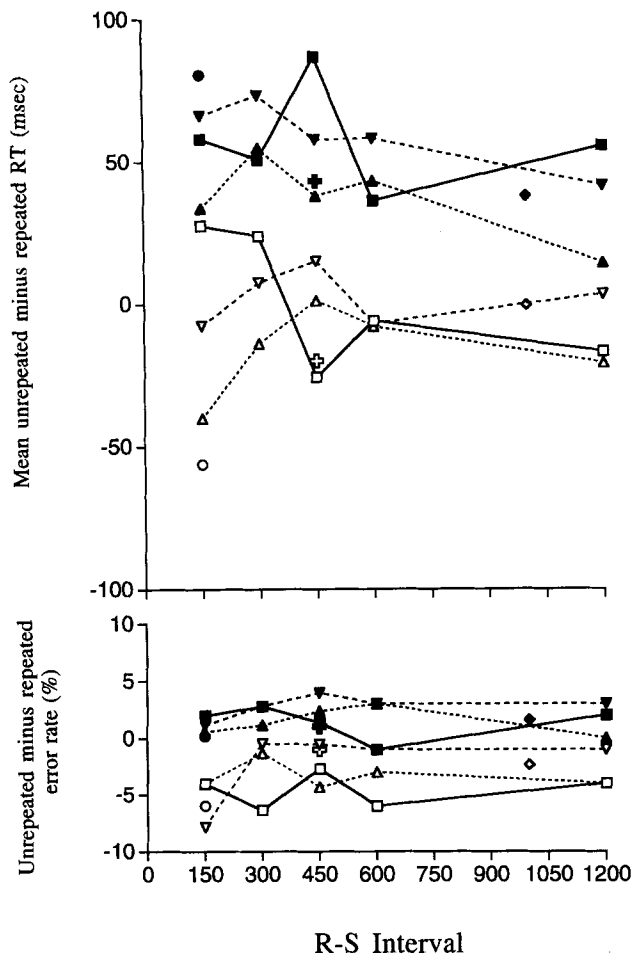
Figure 6. The benefit of a response repetition (i.e., mean reaction time [RT] or error rate for alternated response minus mean RT or error rate for repeated response) within a run (filled symbols) and across a task-switch (open symbols), plotted as a function of response–stimulus (R–S) interval. Experiment 1: circles; Experiment 2: squares; Experiments 3: upright triangles; Experiment 4: inverted triangles; Experiment 5: diamonds; Experiment 6: crosses. In Experiment 2 the interval varied randomly; in the others it was constant for a block or for the whole experiment.

for in executing the new task, a greater increment of response activation is required to overcome the inhibition than would otherwise be needed.

3. It is logically necessary that there exist some mechanism to prevent perseverative reexecution of a response when the stimulus conditions for a response remain present (in sensory memory, if not on the screen). We might assume that, upon execution of a response, its response category is inhibited. Or, there might be a process that checks the relation between a planned response and the previous one to guard against accidental reemission of a response; detection of an imminent repetition could trigger extra stimulus analysis or rechecking of the decision. In a run of trials on a particular task this cost to a repetition is normally more than compensated for by the facilitative effects of repeated exercise of the same category–

response association. However, the switch trials in the present experiments reveal the inhibition/checking cost of a repeated response because switch trials require activation of a different category–response mapping and do not enjoy the benefits of this facilitation.

These hypotheses are clearly speculative, but they raise interesting issues not only about the transfer of control of effector output between tasks but also about the control and monitoring of motor output within tasks. A phenomenon as robust as this interaction between response repetition and task-switching must be telling us something interesting about control processes!

Are Switch Costs Symmetrical?

One feature of Allport et al.'s (1994) study we have not mentioned hitherto is the marked asymmetry of switch cost they reported for their Experiment 5.[4] They found a large switch cost when the participant had to switch from color naming to word naming, but no cost when the participant had to switch from word naming to color naming. This seems counterintuitive. Word naming is the dominant task of the pair, and one might have expected it to be easier to switch to rather than from the dominant task of the pair. Allport et al. argued that this is because a relatively non-dominant S–R mapping requires a more strongly imposed task-set, and this results in greater TSI on subsequent trials when the participant is required to perform a different task.

In the present study, we chose our letter and digit tasks so that they were similar in difficulty. To spare the reader, we have refrained from reporting every test of interaction with task, but Table 8 lists in summary form the baseline performance on both tasks in each experiment, together with the associated time and error switch costs. The interaction between task and switching was significant in only two experiments: The time cost associated with the letter task was significantly smaller than that for the digit task in Experiment 3, $F(1, 9) = 5.5, p < .05, MSE = 48,517$ (but the error cost differed in the opposite direction), and the error cost for the digit task was significantly smaller than that for the letter task in Experiment 4, $F(1, 9) = 3.6, p = .09, MSE = 0.059$. When both RT and accuracy are taken into account, there is little evidence either for a consistent asymmetry of switch costs or for a consistent difference in difficulty between the two tasks. Hence these data do not provide a critical test for Allport et al.'s (1994) conjecture. However, in another experiment (Monsell, Williams, Wright, & Rogers, 1995), we have found that by manipulating the legibility of the stimulus words, it is possible to equate color naming and word naming latencies while preserving the asymmetry of Stroop interference. Under these conditions, switching costs were not asymmetric.

[4] Note that it is not possible to assess this asymmetry in their other experiments, because the costs of switching in each direction were not separable.

ROBERT D. ROGERS AND STEPHEN MONSELL

Table 8

Nonswitch Performance and Switch Costs for Both the Letter and Digit Tasks for Experiments 1 to 6

	Letter				Digit			
	Nonswitch trials		Switch cost		Nonswitch trials		Switch cost	
Experiment	RT	% error	RT	% error	RT	% error	RT	% error
1	679.9	2.8	208.2	4.8	653.1	2.3	241.0	3.7
2	731.2	3.4	190.3	2.7	709.4	3.4	203.9	3.5
3	637.5	3.0	122.6	3.5	619.3	4.1	182.1	3.2
4	475.1	3.7	61.7	3.8	479.2	5.3	49.0	1.6
5	477.0	3.8	57.2	1.7	483.5	3.8	64.8	2.9
6[a]	582.7	4.1	199.4	2.8	583.6	5.3	239.8	3.9

Note. RT = mean reaction time in milliseconds.

[a] Nonswitch performance for Experiment 6 is averaged over Trials 2–4 of a run—see text.

General Discussion

The experiments reported here have shown, like the studies of Jersild (1927), Spector and Biederman (1976), and Allport et al. (1994), that switching predictably between two simple cognitive tasks results in large and reliable increases in RT and error rate. The alternating runs paradigm that we have introduced, in which switch and nonswitch trials are compared within a block, has a major advantage over the technique, invented by Jersild, of comparing performance in blocks of trials alternating between the tasks with blocks of trials on just one task. The switch costs observed within a mixed block of trials in the alternating runs paradigm can be unambiguously attributed to the need to change tasks, not to the requirement to keep two tasks available at the same time or to between-block differences in response criterion, arousal, and so forth.

The switch costs we have observed are clearly not transient phenomena characteristic only of tasks with which the participant is unfamiliar or of an early phase of learning to cope with the requirement to switch tasks. Our participants were given extensive training with each categorization on its own before being introduced to the alternating runs procedure. Extended practice at switching only modestly attenuated, and certainly did not abolish, the switch costs observed, especially if time cost was expressed as a proportion of baseline. In another experiment, Rogers (1993) examined unpredictable switching between semantic and phonological classifications of word stimuli over 4 days of practice; robust switch costs were still present on the 4th day. Thus these switch costs cannot be attributed to executive functions that mediate early learning about individual mappings between stimuli and responses or about switching between them; that is, the switch costs did not occur while certain aspects of task processing were "controlled," only to disappear when they became sufficiently "automatic" (Shiffrin & Schneider, 1977). Of course, very extensive practice might eliminate switch costs, but, if it did, we would suspect that the participant had learned direct associations between each stimulus + task cue combination and the response— *that is, converted the two experimenter-defined S–R mappings into a single S–R mapping.* Our transfer test after

Experiment 1 provided no evidence that this had happened after 2 days of testing.

The switch costs obtained in the alternating runs paradigm exhibit the following properties:

- Allowing participants about half a second to prepare for a predictable task-switch attenuated the time cost of a switch by about one third and also reduced the error cost. We attribute this transient element of switch cost to a stagelike (but incomplete) process of task-set reconfiguration, which is endogenously controlled and can be carried out in anticipation of the stimulus. Its maximum duration is apparently on the order of half a second.

- For longer R–S intervals, there remained a substantial residual switch cost, undiminished as the time available for preparation was increased to over 1 s.

- With a run of four trials on the same task, there was no hint of a graded improvement in performance over the three trials following the trial on which the task-switch was required. The cost of a predictable task-switch seems to be borne entirely by performance on the first trial of a run.

- When the stimulus contained attributes associated with the competing task, both performance on nonswitch trials and the efficiency of switching were impaired. We interpret this as a manifestation of the tendency of stimuli to activate task-sets associated with them—a task-cuing effect. We suppose that this is the same tendency which, when unopposed by endogenous control processes, results in phenomena such as utilization behavior in frontal lobe patients and capture errors under conditions of ordinary absentmindedness.

- Even when the task-switching was successful in the sense of the correct response being elicited, the suppression of the irrelevant S–R mapping was incomplete. Even on the nonswitch trials, but more so on the switch trials, we observed an advantage when the irrelevant attribute was linked, via the S–R mapping specified by the currently irrelevant task, with the same response specified by the relevant attribute—a Stroop-like crosstalk effect.

- However, the crosstalk effect was, at least for RTs, small relative to the task-cuing effect. That is why the latter effect cannot be interpreted purely in terms of competition or reinforcement between multiple responses associated with the stimulus: We need the assumption that the whole task-set is evoked by the associated stimulus attribute. (The contrast between these assumptions can be illustrated by

recent connectionist models of the Stroop effect. In the model of Cohen, Dunbar, and McClelland [1990] "task" nodes and stimulus attribute input nodes are associated with response nodes by feedforward connections via hidden units. There is no way for stimulus attributes to evoke task-sets as such. The task-cuing effect we report here appears to require, at a minimum, the addition of connections that allow stimulus attributes to excite task nodes. In the interactive network presented by Cohen and Huston [1994], by contrast, task and stimulus nodes are bidirectionally connected, so that stimulus attributes can [in principle] activate the associated task as a whole. However, our reference to the latter network is merely illustrative of this one point; we are not claiming that this network can account even for the relative magnitude of the task-cuing and crosstalk effects, let alone the other features of our data.)

- The task-cuing and crosstalk effects of the irrelevant character on switch costs were observed at all R–S intervals and showed no consistent pattern of interaction with R–S interval. Hence they appear to be associated with the residual component of switch cost.

- The response priming effect usually observed when two successive responses happen to be the same was abolished, and in terms of error rate reversed, when a task-switch intervened between the responses. At the very least this suggests that response repetition priming cannot be a property of the programming or execution of the response category alone (cf. Pashler & Baylis, 1991): It is task-specific. These effects were somewhat more marked at a shorter and more predictable R–S interval. We have discussed possible mechanisms above.

The picture of task control suggested by these properties of switch cost is consistent in a general way with Norman and Shallice's production-system metaphor (Norman & Shallice, 1986; Shallice, 1988, 1994). That is, activation of competing action systems, schemas, or (as we prefer) task-sets, triggered by the presence in the environment of certain stimuli, is modulated by an internally initiated control process responsible for ensuring that, at any one time, the appropriate task-set will win and the correct action will be selected. However, substantial elaboration of the metaphor is necessary to account for the details.

The partial reduction of switch costs with an increase in R–S interval is compatible with an endogenously controlled, time-consuming, stagelike process of reconfiguration that can, when opportunity allows, be carried out in anticipation of the stimulus. But the metaphor of this process as operating a literal switch (or set of switches) to connect and disconnect the mappings between stimulus categories and responses is plainly inadequate. The effects of recency of activation of a competing task (Allport et al., 1994) and the effects of the presence in the stimulus of attributes associated with that task (Experiments 1, 2, 3, and 6 of the present article) are both congruent with a poststimulus competition between task-sets whose outcome is merely biased by any prestimulus reconfiguration. Even when participants are as prepared for a stimulus as they can be, S–R associations mediating a competing task afforded by the same stimuli remain connected and, if recently activated by performance of that task or concurrently activated by irrelevant attributes of the stimulus, are a significant source of

competition with the S–R associations mediating the required task, prolonging "the time needed for the system to settle to a unique response-decision," as Allport et al. (1994, p. 437) put it.

TSI does two kinds of work in Allport et al.'s theory. It accounts for the fact that switching to a task is harder when another task afforded by the same stimulus has been exercised in the last few minutes (Allport et al., 1994, Experiment 4): To do this work, TSI must persist for a few minutes. It is also assumed to account entirely for the residual component of switch cost, from which the opportunity to engage in endogenous activity prior to the stimulus provides no escape, as in Allport et al.'s Experiment 5 and our Experiments 2 through 6. If TSI is as persistent as Allport et al. proposed, then it should retard performance not just on the first trial after a task-switch, but on many thereafter. We have not shown that this does not happen. What we have shown, in Experiment 6, is that there is a substantial component of switch cost that is borne only by the first trial of a run of trials on a given task, with no further improvement in performance between the second and fourth trials of a run. To this extent at least, the metaphor of a switchlike process does characterize task-set reconfiguration. We conclude that there is a component of the process of switching tasks that cannot be executed in anticipation of the stimulus but needs to be executed after only the first of several stimuli requiring performance of that task. We propose that this is the completion of a stagelike process of reconfiguration and that completion can be triggered only exogenously by the arrival of a stimulus suitably associated with the task.

Finally, we return to the initial decline in the switch cost as the R–S interval was extended up to about half a second (as robustly demonstrated in our Experiments 3 and 4 and rather less robustly in Experiment 5 of Allport et al., 1994). Whereas we interpret this decline as being due to the participant's engaging in anticipatory task-set reconfiguration when time permits, Allport et al. (1994) dismissed it as merely a short-term component of the dissipation of TSI by analogy to the transient component of proactive interference effects reported in the memory literature (e.g., Loess & Waugh, 1967). For the TSI theory to work at all, however, the dissipation of TSI must be an entirely passive process; otherwise participants could actively suppress TSI and avoid switch costs at longer intervals. But why, then, was there no decline in the switch cost with variable R–S intervals in Experiment 2? If the dissipation of TSI is a passive process, and the effects of TSI are sufficient to account for switch cost, then switch cost should dissipate through the R–S interval at the same rate whether the interval is varied from trial to trial or constant. The decline in switch cost with R–S interval when the interval is held constant (Experiments 3 and 4), coupled with the complete absence of such a decline when the interval is unpredictable (Experiment 2), suggests, in contrast, that the decline in switch cost is due to an active process. Participants engaged in this active process when the interval could be predicted, but not when it was unpredictable. We propose that this active process is the endogenous component of task-set configuration.

References

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

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.

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

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

Bertelson, P. (1967). The time course of preparation. *Quarterly Journal of Experimental Psychology, 19,* 272–279.

Bertelson, P., & Boons, J. P. (1960). Time uncertainty and choice reaction time. *Nature, 187,* 531–532.

Bertelson, P., & Tisseyre, F. (1968). The time course of preparation with regular and irregular foreperiods. *Quarterly Journal of Experimental Psychology, 20,* 297–300.

Bertelson, P., & Tisseyre, F. (1969). The time course of preparation: Confirmatory results with visual and auditory warning signals. *Acta Psychologica, 30,* 145–154.

Biederman, I. (1973). Mental set and mental arithmetic. *Memory and Cognition, 1,* 383–386.

Briand, K. A., & Klein, R. M. (1987). Is Posner's "beam" the same as Treisman's "glue"? On the relation between visual orienting and feature integration. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 228–241.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332–361.

Cohen, J. D., & Huston, T. A. (1994). Progress in the use of interactive models for understanding attention and performance. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 453–476). Cambridge, MA: MIT Press.

Dixon, P., & Just, M. A. (1986). A chronometric analysis of strategy preparation in choice reactions. *Memory and Cognition, 14,* 488–500.

Drewe, E. A. (1974). The effect of type and area of brain lesion on Wisconsin Card Sorting Test performance. *Cortex, 10,* 159–170.

Duncan, J. (1975). Response selection rules in spatial choice reaction tasks. In S. Dornic (Ed.), *Attention and performance VI* (pp. 49–61). Hillsdale, NJ: Erlbaum.

Duncan, J. (1979). Divided attention: The whole is more than the sum of its parts. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 216–228.

Frith, C. D., & Done, J. D. (1986). Routes to action in reaction time tasks. *Psychological Research, 48,* 169–177.

Gottsdanker, R. (1975). The attaining and maintaining of preparation. In P. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 33–42). London: Academic Press.

James, W. (1890). *The principles of psychology*. New York: Holt.

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

Klemmer, E. T. (1956). Time uncertainty in simple reaction time. *Journal of Experimental Psychology, 31,* 179–184.

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

Loess, H., & Waugh, N. C. (1967). Short-term memory and inter-

trial interval. *Journal of Verbal Learning and Verbal Behaviour, 6,* 455–460.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

Milner, B. (1963). Some effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology, 9,* 90–100.

Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General, 118,* 43–71.

Monsell, S., Taylor, T., & Murphy, K. (1995). *Stroop interference as task-set interference: Evidence from effects of lexicality and word frequency*. Manuscript in preparation.

Monsell, S., Williams, N., Wright, I., & Rogers, R. D. (1995). *Asymmetry of Stroop interference and the costs of switching between word naming and colour naming*. Manuscript in preparation.

Näätänen, R. (1970). The diminishing time uncertainty with the lapse of time after the warning signal in reaction-time experiments with varying foreperiods. *Acta Psychologica, 34,* 399–419.

Näätänen, R., & Merisalo, A. (1975). Expectancy and preparation in simple reaction time. In S. Dornic (Ed.), *Attention and performance VI* (pp. 115–138). Hillsdale, NJ: Erlbaum.

Navon, D., & Miller, J. (1987). Role of outcome conflict in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 435–448.

Nelson, H. E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex, 12,* 313–324.

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

Pashler, H., & Baylis, G. (1991). Procedural learning: Intertrial repetition effects in speeded-choice tasks. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 33–48.

Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology, 41A*(1), 19–45.

Posner, M. I., Klein, R., Summers, J., & Buggie, S. (1973). On the selection of signals. *Memory and Cognition, 1,* 2–12.

Rabbitt, P. M. A. (1968). Repetition effects and signal classification strategies in serial choice-response tasks. *Quarterly Journal of Experimental Psychology, 20,* 232–240.

Rabbitt, P. M. A., & Vyas, S. (1970). An elementary preliminary taxonomy for some errors in laboratory choice RT tasks. *Acta Psychologica, 33,* 56–76.

Rabbitt, P. M. A., & Vyas, S. (1973). What is repeated in the "repetition effect"? In S. Kornblum (Ed.), *Attention and performance IV* (pp. 327–342). New York: Academic Press.

Reason, J. T. (1984). Lapses of attention. In R. Parasuraman, R. Davies, & J. Beatty (Eds.), *Varieties of attention* (pp. 515–549). Orlando, FL: Academic Press.

Rogers, R. D. (1993). *The costs of switching between cognitive tasks: A performance analysis*. Unpublished doctoral dissertation, University of Cambridge, Cambridge, England.

Shaffer, L. H. (1965). Choice reaction with variable mapping. *Journal of Experimental Psychology, 70,* 284–288.

Shaffer, L. H. (1967). Transition effects in three-choice reaction with variable S–R mapping. *Journal of Experimental Psychology, 73,* 101–108.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge, England: Cambridge University Press.

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

Shallice, T., Burgess, P. W., Schon, R., & Baxter, D. M. (1989). The origins of utilization behaviour. *Brain, 112,* 1587–1598.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

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

Sudevan, P., & Taylor, D. A. (1987). The cuing and priming of cognitive operations. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 89–103.

Thomas, E. A. C. (1973). On a class of additive learning models: Error correcting and probability matching. *Journal of Mathematical Psychology, 10,* 241–264.

Thomas, E. A. C., & Legge, D. (1970). Probability matching as a basis for detection and recognition decisions. *Psychological Review, 77,* 65–72.

Research Awards in Experimental Psychology

The Division of Experimental Psychology of the American Psychological Association (Division 3) announces a series of up to five annual research awards. These reviews are to be based on review of the research submitted to or published in the APA's Journals of Experimental Psychology each year by relatively new investigators. The intention is to provide early recognition to new scholars whose research contributions are especially promising. These awards are

Division of Experimental Psychology (Annual)
Young Investigator Award in Experimental Psychology:
Animal Behavior Processes;

Division of Experimental Psychology (Annual)
Young Investigator Award in Experimental Psychology:
Human Perception and Performance;

Division of Experimental Psychology (Annual)
Young Investigator Award in Experimental Psychology:
Learning, Memory, and Cognition;

Division of Experimental Psychology (Annual)
Young Investigator Award in Experimental Psychology:
General;

and

Division of Experimental Psychology (Annual)
Young Investigator Award in Experimental Psychology:
Applied.

These awards will be announced and given to the winners each year at the Division 3's business meeting held at the APA annual convention.