# Concurrent Response-Selection Processes in Dual-Task Performance: Evidence for Adaptive Executive Control of Task Scheduling

Eric H. Schumacher
University of Michigan

Erick J. Lauber
University of Georgia

Jennifer M. Glass, Eileen L. Zurbriggen, Leon Gmeindl, David E. Kieras, and David E. Meyer
University of Michigan

This article reports 4 experiments that used the psychological refractory period procedure to characterize how people perform multiple tasks concurrently. For each experiment, a primary choice-reaction task was paired with a secondary choice-reaction task that had two levels of response-selection difficulty. Experiments 1 and 2 varied secondary-task response-selection difficulty by manipulating the number of stimulus–response (S-R) pairs. The effect of this factor on secondary-task reaction times (RTs) decreased reliably as the stimulus onset asynchrony (SOA) decreased. Experiments 3 and 4 varied secondary-task response-selection difficulty by manipulating S-R compatibility. Again, the effect of this factor on secondary-task RTs decreased reliably as SOA decreased. These results raise doubts about the existence of an immutable structural response-selection bottleneck and suggest that response selection for 2 concurrent tasks may overlap temporally.

The study of human multiple-task performance has a long history in cognitive psychology (for reviews, see Meyer & Kieras, 1997a, 1997b; Pashler, 1994a). Research in this area is important practically because it may produce results of benefit to people who must perform multiple tasks consistently well under real-world circumstances (e.g., air-traffic controllers and aircraft pilots). Moreover, research in this area is important theoretically because the concurrent performance of multiple tasks imposes heavy demands on the human information-processing system and thus potentially enables deep insights into how the system's components are functionally organized and implemented.

One experimental procedure that has been especially

useful for investigating the nature of human dual-task performance is the *psychological refractory period* (PRP) procedure. In the PRP procedure, there is a series of discrete trials, and participants must respond to two successively presented stimuli on each trial. The time between the first and second stimuli of a trial is the stimulus onset asynchrony (SOA). Participants are typically instructed to respond quickly and accurately to both stimuli but to give primary emphasis to the first one. As a result, reaction times (RTs) for the first stimulus are usually little affected by the SOA, whereas RTs for the second stimulus often increase substantially as the SOA decreases, yielding a *PRP effect.*

To account for the PRP effect, theorists have proposed many explanations. Perhaps the most prominent of these is the *response-selection bottleneck* (RSB) hypothesis (Pashler, 1994a; Welford, 1980). According to this hypothesis, performing each task in the PRP procedure requires a series of processing stages, among which is a response-selection stage that can deal with only one task at a time. Because of this putative limited capacity, the RSB hypothesis has an important implication: When the SOA is short and the response-selection stage is devoted to the primary task, response selection for the secondary task must wait temporarily, yielding the PRP effect.

To test the RSB hypothesis, some researchers have manipulated secondary-task factors that affect various processing stages and have used *locus-of-slack logic* to interpret their results (Carrier & Pashler, 1995; De Jong, 1993; Fagot & Pashler, 1992; McCann & Johnston, 1992; Pashler, 1984; Pashler & Johnston, 1989; Ruthruff, Miller, & Lachmann, 1995; Van Selst & Jolicoeur, 1997). For example, manipulations of SOA and Task 2 stimulus intensity, which affects secondary-task stimulus encoding, have been found to produce underadditive effects on mean Task 2 RTs (De Jong, 1993; Pashler, 1984; Pashler & Johnston, 1989). In contrast,

manipulations of SOA and Task 2 factors such as stimulus repetition (Pashler & Johnston, 1989), memory-retrieval difficulty (Carrier & Pashler, 1995), target presence–absence (Pashler, 1984), mental-rotation angle (Ruthruff et al., 1995), stimulus–response (S-R) numerosity (Van Selst & Jolicoeur, 1997), and S-R compatibility (Fagot & Pashler, 1992; McCann & Johnston, 1992)—all of which affect secondary-task response selection—have been found to produce additive effects on mean Task 2 RTs. As Figure 1 shows, this pattern of factor effects could stem from a "bottleneck" in response selection of the sort proposed under the RSB hypothesis. At short SOAs, factor effects on Task 2 processing stages before, but not after, the bottleneck would be absorbed in the processing slack that occurs while performance of Task 2 waits for the completion of Task 1 response selection.

Yet the exact nature of the RSB remains unclear. Pashler (1984, 1994a, 1994b) and others (e.g., De Jong, 1993; McCann & Johnston, 1992; Pashler & Johnston, 1989; Welford, 1980) have proposed that it is an immutable structural cognitive mechanism. According to this proposal, the human brain is "wired" such that it cannot select the response to more than one stimulus at a time. If so, then

manipulations of SOA and Task 2 response-selection difficulty should always affect mean Task 2 RTs additively during the PRP procedure. However, some studies that used this procedure yielded underadditive effects of these factors (De Jong, 1993; Hawkins, Rodriquez, & Reicher, 1979; Ivry, Franz, Kingstone, & Johnston, 1998; Karlin & Kestenbaum, 1968; Meyer et al., 1995). For example, Figure 2 shows mean Task 2 RTs from studies by Karlin and Kestenbaum (1968) and Hawkins et al. (1979) in which Task 2 response-selection difficulty was manipulated by varying the number of alternative S-R pairs. These results suggest that in at least some situations, a bottleneck may occur after rather than before the Task 2 response-selection stage, and people may select the responses for two tasks concurrently, contrary to the RSB hypothesis (Keele, 1973; Meyer & Kieras, 1997a, 1997b, 1999; Meyer et al., 1995).

Given these and other complementary considerations, Meyer and Kieras (1997a, 1997b, 1999; Meyer et al., 1995) formulated an alternative account of dual-task performance for the PRP procedure. They proposed a class of *adaptive executive control* (AEC) models in which it is assumed that people have flexible control over the course of secondary-task processing stages. Under AEC models, such control is
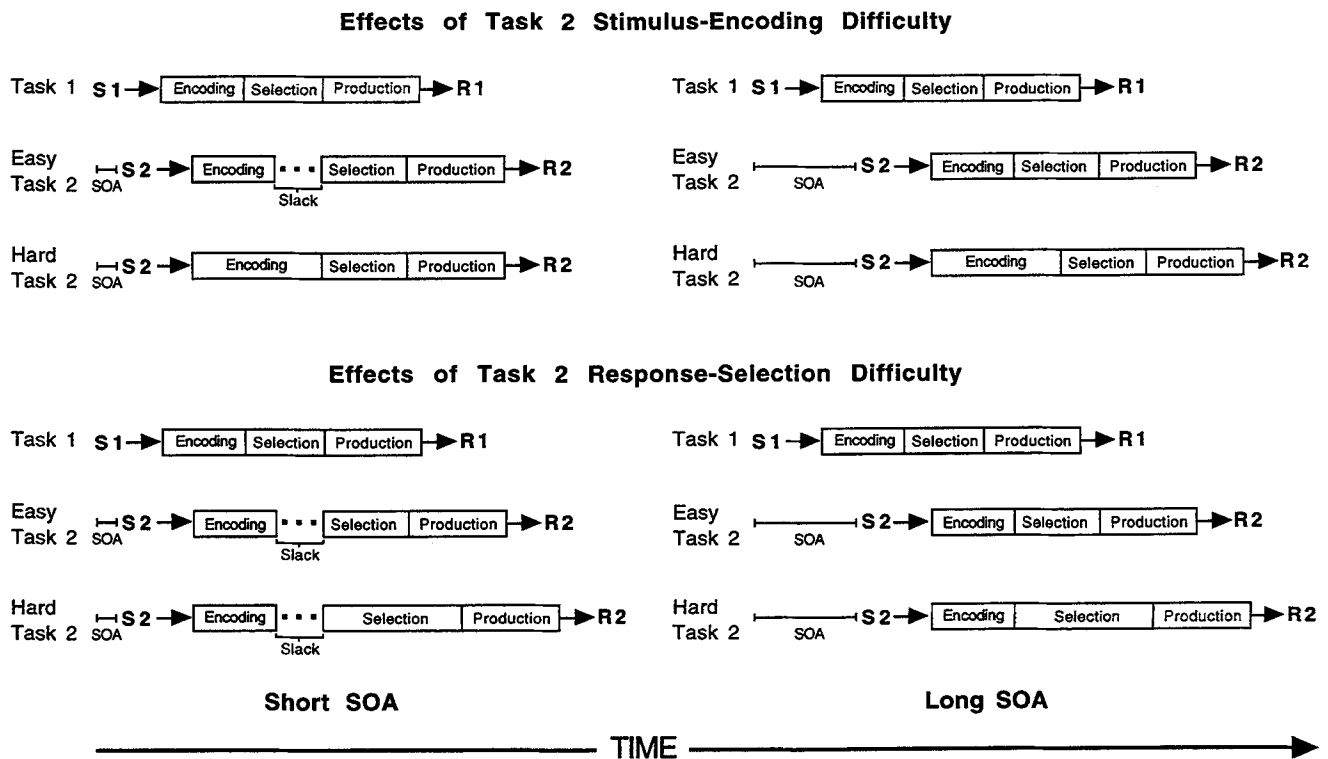


*Figure 1.* Stages of processing with a response-selection bottleneck for dual-task performance in the psychological refractory period (PRP) procedure. Here processing for Task 1 and Task 2 begins with the presentation of a task stimulus (S), continues through stimulus encoding, response selection, and movement production, and ends with the output of an overt response (R). The upper half of the figure shows how manipulations of Task 2 stimulus-encoding difficulty and stimulus onset asynchrony (SOA) produce underadditive effects on Task 2 reaction times. The lower half of the figure shows how manipulations of Task 2 response-selection difficulty and SOA produce additive effects on Task 2 reaction times.

*Figure 2.* Mean Task 2 reaction times as a function of Task 2 response-selection difficulty and stimulus onset asynchrony (SOA) in psychological refractory period studies by (A) Karlin and Kestenbaum (1968) and (B) Hawkins, Rodriquez, and Reicher (1979).
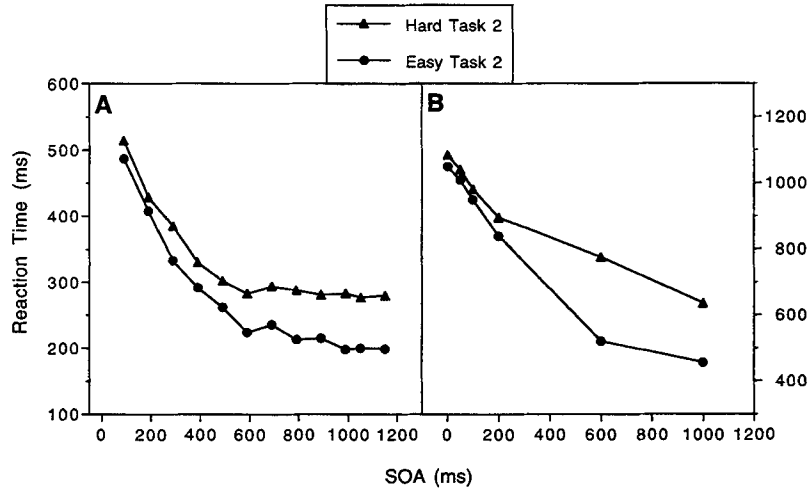
achieved through executive processes that can lock out (suspend) and unlock (resume) performance of Task 2 between any two processing stages (see Figure 3).

According to this view, the *Task 2 lockout point* is a stage of processing for Task 2 such that when it is about to start, further progress on Task 2 is suspended temporarily until Task 1 is deemed to be completed. The *Task 1 unlocking event* is a stage of processing for Task 1 such that when it has

ended, Task 1 is deemed to be completed. When the Task 1 unlocking event occurs, executive processes unlock Task 2 and let the secondary-task stages proceed to completion from the point at which they were previously suspended. Like the choice of a decision criterion (beta) in signal-detection theory (Tanner & Swets, 1954), the specifications of particular Task 2 lockout points and Task 1 unlocking events by the executive processes are presumably optional.



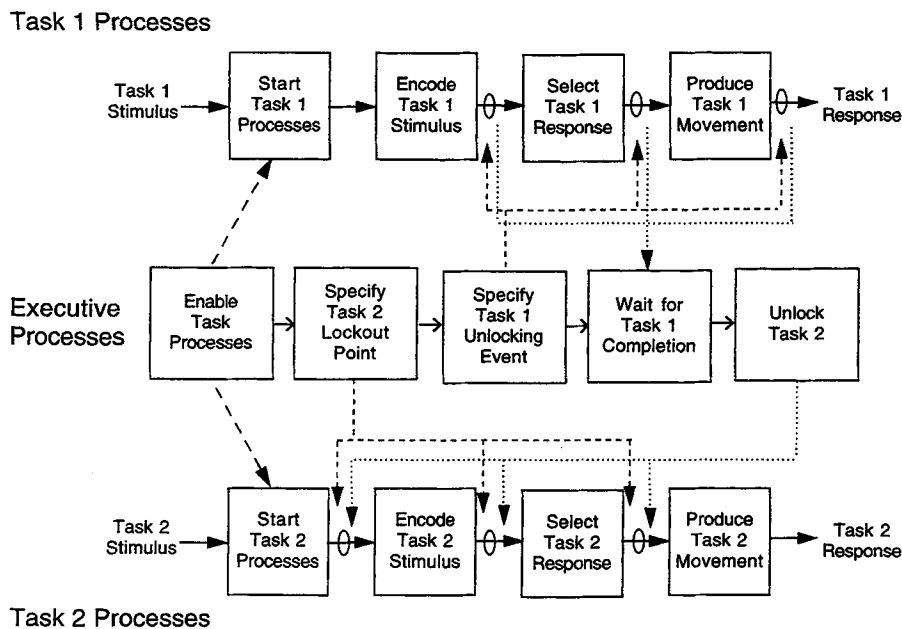*Figure 3.* Component processes for adaptive executive control models whereby tasks in the psychological refractory period procedure may be flexibly scheduled. Arrows leading to and from various executive processes denote where a Task 2 lockout point may be set, where a Task 1 unlocking event may occur, where processing for Task 1 is deemed to be completed, and where processing for Task 2 should be resumed.

With these specifications, executive processes may implement various alternative "software" bottlenecks and scheduling strategies, depending on relative task priorities, participants' strategic biases, participants' amount of practice, or other ancillary factors. The specific locations of the lockout point in Task 2 and of the unlocking event in Task 1 yield particular subtypes of AEC models.

For example, under some conditions, extreme emphasis on Task 1 in the PRP procedure may lead participants to adopt a cautious scheduling strategy, producing a strategic rather than structural RSB. To implement such cautious scheduling, executive processes would lock out Task 2 before secondary-task response selection and unlock it after completion of primary-task response selection. Under other conditions, however, participants may adopt a more daring scheduling strategy, allowing response-selection processes for the two tasks to overlap temporally by not locking out Task 2 until after secondary-task response selection has been completed. The potential for such flexible task scheduling is suggested by the diverse patterns of additivities and interactions that previous investigators have found between the effects of SOA and of Task 2 response-selection difficulty.

Computer simulations using the heuristic task-scheduling principles of our AEC models have successfully reproduced the results from PRP experiments by Karlin and Kestenbaum (1968), Hawkins et al. (1979), and McCann and Johnston (1992), as well as the results from other relevant dual-task studies (e.g., Ballas, Heitmeyer, & Perez, 1992). The results of these simulations suggest that people may not have an immutable structural RSB (Meyer & Kieras, 1997a, 1997b).

Here we report four experiments that provide further evidence that RSBs are not structural and immutable. Instead, our experiments show that these bottlenecks may be optional and strategic. Our experiments also examine the effect of practice on participants' choice of task-scheduling strategies, which is an important aspect of dual-task performance that has not yet received sufficient investigation (but see Gopher, 1993; Van Selst, Ruthruff, & Johnston, 1997).

In Experiment 1 we replicated and extended prior work by Karlin and Kestenbaum (1968) using S-R numerosity to manipulate response-selection difficulty in an auditory-manual secondary task. Like Karlin and Kestenbaum, we found an underadditive interaction between the effects of this factor and SOA on mean Task 2 RTs. The substantial temporal overlap between response-selection processes for Tasks 1 and 2 found in Experiment 1 casts further doubt on the RSB hypothesis.

To decrease the likelihood that this response-selection manipulation also affected the difficulty of stimulus encoding, in Experiment 2 we replicated and extended prior work by Hawkins et al. (1979) using symbolic stimuli (viz., digits) instead of tones for the Task 2 stimuli. Again, like Hawkins et al., we found an underadditive interaction between the effects of this factor and SOA on mean Task 2 RTs. This outcome further demonstrates the occurrence of temporal overlap between Task 1 and Task 2 response-selection processes.

In Experiment 3, which replicates and extends prior work by McCann and Johnston (1992), we manipulated another popular response-selection factor, S-R compatibility. The results of Experiment 3 reveal that participants' task-scheduling strategies may change with practice. Early in practice, participants showed evidence of a cautious scheduling strategy with a strategic RSB. After practice, however, they came to use a daring scheduling strategy that overlaps response-selection processes for the primary and secondary tasks.

Finally, in Experiment 4 we replicated Experiment 3 using separate response modalities for the two tasks. The results from Experiment 4 disarm the possibility that the previous underadditive interactions stemmed from a secondary immutable structural bottleneck in movement production. Taken together, our four experiments therefore provide substantial new support for the AEC models of Meyer and Kieras (1997a, 1997b, 1999).

## Experiment 1

Task 1 of this experiment was a two-choice visual-manual task. Task 2 was an auditory-manual task in which the correct responses were specified by the frequencies of presented tones. We used an auditory-manual secondary task because encoding the stimuli for it did not require eye movements, which can preclude concurrent response-selection processes (Meyer & Kieras, 1997a, 1997b). Task 2 response-selection difficulty was manipulated by varying S-R numerosity. This manipulation is appropriate because prior research has shown that S-R numerosity mainly affects response selection (Brainard, Irby, Fitts, & Alluisi, 1962; Broadbent & Gregory, 1965; Morin & Forrin, 1965; Sanders, 1980; Sternberg, 1969; Theios, 1973).

In prior research, where each alternate stimulus was mapped onto a different response, S-R numerosity typically has been manipulated by increasing both the number of stimuli and the number of responses. However, in Experiment 1, the number of responses was held constant across the levels of S-R numerosity for Task 2. We did this to replicate prior conditions in the experiment by Karlin and Kestenbaum (1968) and to preclude undesirable ancillary effects of response numerosity on movement production (Miller & Ulrich, 1998). Despite the latter constraint, it nevertheless seems plausible that our manipulation of S-R numerosity would have effects on response-selection difficulty similar to those of past S-R numerosity manipulations, as suggested by the results of Karlin and Kestenbaum.[1]

As stated previously, Karlin and Kestenbaum (1968) found a large underadditive interaction between the effects of S-R numerosity and SOA on mean Task 2 RTs in a PRP procedure with auditory-manual tasks similar to ours. Their S-R numerosity effect was 81 ms at the longest SOA but

---

[1] Furthermore, we assume that our particular manipulation of S-R numerosity affected stimulus-encoding processes no more than other such manipulations have in the past. Because previous effects of S-R numerosity on stimulus encoding appear to have been relatively small (Brainard et al., 1962; Morin & Forrin, 1965; Sanders, 1980; Sternberg, 1969; Theios, 1973), the present ones were presumably also quite small.

only 27 ms at the shortest SOA (see Figure 2A). By locus-of-slack logic, this underadditivity implies that the response-selection processes for Tasks 1 and 2 were concurrent.

However, Karlin and Kestenbaum's (1968) easy and hard levels of response-selection difficulty involved simple-reaction and choice-reaction tasks, respectively. This complicates the interpretation of their results. Pashler (1994a) argued that RTs for simple-reaction and choice-reaction tasks cannot validly be compared because "PRP effects observed in simple RT seem fundamentally different from those found in choice RT" (p. 229). It is not clear to us exactly how these fundamental differences obviate the results of Karlin and Kestenbaum; nevertheless we take Pashler's concern seriously. Therefore, to disarm it, we made both levels of Task 2 response-selection difficulty in Experiment 1 involve choice-reaction tasks.

Another related objective of Experiment 1 was to assess claims made by Van Selst and Jolicoeur (1997) that the pattern of results reported by Karlin and Kestenbaum (1968) could not be replicated. Contrary to the results of Karlin and Kestenbaum, Van Selst and Jolicoeur failed to find underadditive effects of SOA and S-R numerosity on mean Task 2 RTs. However, the results of our first experiment show that such interactions are obtainable under conditions that replicate and extend those of Karlin and Kestenbaum.

## Method

### Participants

Eleven right-handed undergraduate students (5 men and 6 women) at the University of Michigan participated as paid volunteers. They had normal or corrected-to-normal vision and were paid $5.00 per hour plus bonuses based on the quality of their performance. The data from 3 participants were not analyzed because 2 of them (1 man and 1 woman) had error rates greater than 15% and the other (1 woman) did not learn the tasks quickly enough (thus, $N = 8$ in the analyses).

### Apparatus

Participants sat about 80 cm from a display screen in a quiet, semi-dark room. Visual stimuli were presented on a Zenith ZVM-1200 monochrome monitor with an AST Premium 386 personal computer. Auditory stimuli were presented over Sennheiser HMD 24 headphones. Responses were made with a piano-type response keyboard. It had two groups of five finger keys, with one group for each hand.

### Design and Procedure

*Tasks.* Participants performed two tasks during the experiment. Task 1 was a visual-manual task. On each Task 1 trial, either a 2 or a 3 appeared in the center of the display monitor. Participants responded by pressing the left index-finger key for the 2 or the left middle-finger key for the 3. Task 2 was an auditory-manual task. On each Task 2 trial, one of four tones (330, 500, 1120, or 1650 Hz) was presented. If the 500-Hz or 1120-Hz tone occurred, participants had to press the right index-finger key. If the 330-Hz or 1650-Hz tone occurred, participants had to press the right middle-finger key. Task 2 had two (easy and hard) levels of response-

selection difficulty. The hard Task 2 used all four possible S-R pairs, whereas the easy Task 2 used only two S-R pairs (viz., those with the 1120-Hz and 1650-Hz tones). The stimuli for the two tasks were separated by one of five SOAs: 50, 150, 250, 500, or 1,000 ms.

*Sessions.* Two test sessions were conducted on separate days for each participant. Session 1 included 18 trial blocks. The first 6 of these were single-task blocks, with two blocks for Task 1 and two for each level of Task 2 difficulty. The remainder of the session consisted of 12 dual-task blocks (i.e., Task 1 paired with either the easy or hard Task 2 on each trial). There were 24 trials per single-task block and 40 trials per dual-task block.

Session 2 included 18 dual-task blocks with 40 trials each. The first two of these involved the easy and hard versions of Task 2, respectively, with stimuli separated by a 500-ms SOA on each trial. Subsequent blocks used all five SOAs. The third and fourth of these again involved the easy and hard versions of Task 2, respectively. During the remainder of the session, dual-task blocks were paired such that for each block pair, either the easy or hard Task 2 had to be performed, and the difficulty of Task 2 alternated across block pairs in a counterbalanced fashion. All of Session 1 and the first two blocks of Session 2 were for preliminary warm-up and were omitted from subsequent data analyses.

*Trial blocks.* All possible digit–tone–SOA combinations occurred equally often within each trial block. Participants were told before each block which task(s) would be involved.

*Trials.* The trials of each block began with a fixation cross presented in the center of the display monitor. On dual-task trials, 500 ms after the onset of the fixation cross, a Task 1 digit replaced the cross, and after the SOA, a Task 2 tone sounded for 40 ms. On single-task trials, only one stimulus was presented. The assignments of stimuli and SOAs to particular trials were randomized within each block.

*Feedback and points.* After each trial, participants received point totals that depended on their performance. For both tasks, we awarded 100 points per correct response and deducted 1 point per every 10 ms taken to respond correctly; 100 points were deducted per incorrect response. On dual-task blocks, participants had to make their Task 1 response first or else both responses were considered to be incorrect.[2] In addition, participants received an extra 1,000 points after each dual-task block on which their mean Task 1 RT at the 50-ms SOA was within 75 ms of their mean Task 1 RT at the 1,000-ms SOA. This reward system encouraged participants to complete Task 1 as quickly as possible regardless of the SOA and discouraged grouping of Task 1 and Task 2 responses. Participants earned $1 for every 30,000 points they scored. They were fully informed about the reward system before the experiment began.

After each trial block, participants received more detailed feedback about their numbers of correct responses, mean RTs, and points. Also, during the first six blocks of Session 1 and the first two blocks of Session 2, feedback about response accuracy and points was given after each trial. Subsequently, feedback about response accuracy occurred only after trials on which there were incorrect responses.

## Results

RTs from dual-task trials on which both responses had been correct were analyzed. Outliers in the RT data were

---

[2] Out-of-order errors occurred on fewer than 0.10% of trials overall.

trimmed using a systematic algorithm.[3] This algorithm, which precluded potential distortions by RT outliers, removed 4.5% of the trials in the overall data set, leaving 6,821 response pairs. The remaining RT data were analyzed separately for Task 1 and Task 2 using a within-subjects analysis of variance (ANOVA) with Task 2 response-selection difficulty (i.e., S-R numerosity) and SOA as factors. Figure 4 shows the mean RTs of correct responses for each task as a function of Task 2 response-selection difficulty and SOA.

Task 1 error rates were computed without regard for Task 2 performance and were analyzed through an ANOVA with Task 2 response-selection difficulty and SOA as factors. Task 2 error rates were analyzed separately through a similar ANOVA. Only Task 2 errors that occurred after correct Task 1 responses were analyzed. Table 1 shows the error rates for each task as a function of Task 2 difficulty and SOA. In what follows, we discuss the results for each task.

## Task 1

*Reaction times.* The main effect of Task 2 response-selection difficulty on mean Task 1 RTs was marginally reliable, $F(1, 7) = 5.34$, $.05 < p < .06$. Mean Task 1 RTs were longer when participants performed the hard Task 2 than when they performed the easy Task 2. However, this RT difference was small (less than 10 ms on average). Neither the main effect of SOA, $F(1, 7) = 1.85$, $p > .10$, nor the interaction between the effects of Task 2 difficulty and SOA, $F(4, 28) = 2.23$, $p > .05$, was reliable.

*Error rates.* The overall error rate for Task 1 was 5.44%. The Task 2 difficulty effect was small but reliable, $F(1, 7) = 12.07$, $p < .05$. Participants made more errors on Task 1 when Task 2 was hard (6.88%) than when Task 2 was easy (4.01%). The accompanying SOA effect was not reliable, $F(4, 28) = 2.53$, $p > .05$, but the interaction between the
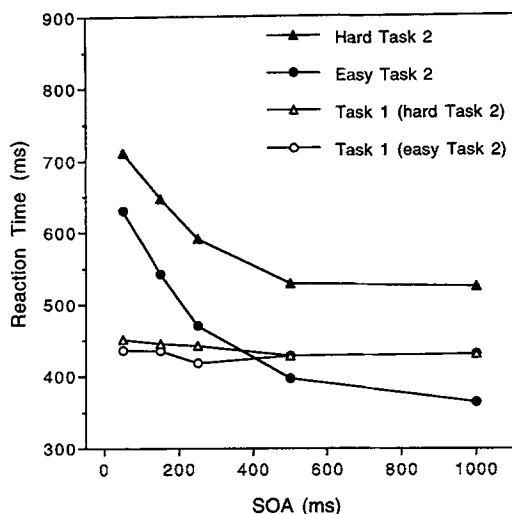


*Figure 4.* Mean reaction times for Tasks 1 and 2 as a function of Task 2 response-selection difficulty and stimulus onset asynchrony (SOA) in Experiment 1.

Table 1

*Mean Error Rates (%) in Experiment 1 for Each Task as a Function of Task 2 Response-Selection Difficulty and Stimulus Onset Asynchrony (SOA)*

| Task 2 type | SOA (ms) | | | | |
|---|---|---|---|---|---|
| | 50 | 150 | 250 | 500 | 1,000 |
| Task 1 | | | | | |
| Hard | 5.85 | 7.06 | 7.81 | 7.04 | 6.64 |
| Easy | 7.99 | 4.50 | 2.70 | 3.09 | 1.75 |
| Task 2 | | | | | |
| Hard | 1.54 | 0.00 | 0.60 | 0.60 | 0.79 |
| Easy | 1.93 | 1.96 | 0.20 | 0.75 | 1.56 |

effects of Task 2 difficulty and SOA was reliable, $F(4, 28) = 4.10$, $p < .01$. As SOA decreased, the number of errors on Task 1 tended to increase when Task 2 was easy but did not change much when Task 2 was hard.

## Task 2

*Reaction times.* For mean Task 2 RTs, main effects of Task 2 response-selection difficulty and SOA, as well as their interaction, were reliable. Overall, mean Task 2 RTs were faster when Task 2 was easy than when Task 2 was hard, $F(1, 7) = 134.96$, $p < .0005$. As the SOA decreased, mean Task 2 RTs increased, $F(4, 28) = 107.45$, $p < .0005$ (see Figure 4). However, the Task 2 difficulty effect decreased as SOA decreased, $F(4, 28) = 5.67$, $p < .005$.[4]

---

[3] Most popular outlier-removal procedures suffer from one or more of several deficiencies: (a) ignoring the mean of the data set, (b) ignoring the distribution of the data set, and (c) using outliers in computing the mean and standard deviation for the data set. We therefore adopted an outlier-removal procedure that overcame these deficiencies. Here the RT data were first sorted into cells defined by particular combinations of the experimental factors. Second, after this sorting, the logarithm of each correct-response RT in each cell of the experimental design was computed. This yielded transformed observations whose distributions more closely approximated Gaussian ones than did those of the untransformed observations. Third, 10% of the logarithmically transformed RTs from each cell's distribution tails were temporarily disregarded (trimmed). This precluded possible outliers from contributing to estimates of the standard deviations for the RT distributions in the design cells. Fourth, the means and standard deviations of the remaining RT logarithms were estimated for each cell. Fifth, each of these standard deviations was multiplied by an adjustment factor of 1.512 to compensate for the tails that had been trimmed. This yielded standard-deviation estimates for the distributions of RT logarithms in the various cells that were presumably unbiased by the presence of outliers. Sixth, RTs whose logarithms differed by more than 3.896 adjusted *SD*s from their respective trimmed means were then removed from each cell, leaving the remaining data that contributed to the reported analyses.

[4] In each of the four experiments, the interactions between the effects of Task 2 response-selection difficulty and SOA on mean Task 2 RTs that were found to be reliable in analyses of the trimmed data set were also reliable in analyses of the original untrimmed data sets.

Figure 5 shows the response-selection difficulty effects on mean Task 2 RTs at the shortest and longest SOAs. The 79-ms difference between the Task 2 difficulty effects at these SOAs was reliable, $t(7) = 3.82, p < .01$.

*Error rates.* The overall Task 2 error rate was low (0.99%). There were no reliable main effects: $F(1, 7) = 2.33, p > .15$, and $F(4, 28) = 2.18, p > .05$, for Task 2 difficulty and SOA, respectively. The interaction between the effects of Task 2 difficulty and SOA was small but reliable, $F(4, 28) = 2.86, p < .05$. SOA had a slightly greater effect on error rates for the easy Task 2 than for the hard Task 2, but the error rates were less than 2% in all cases.

## Discussion

The present underadditive interaction between the effects of Task 2 response-selection difficulty and SOA on mean Task 2 RTs (see Figure 4) provides new evidence against the traditional RSB hypothesis. We found the difficulty effect to be 75 ms less on average at the shortest SOA than at the longest SOA, embodying a 48% decrease (see Figure 5), which suggests that participants' response-selection processes for Tasks 1 and 2 had extensive temporal overlap at the shortest SOA. Consistent with AEC models (see Figure 3), it appears that RSBs are not immutable or structural, but instead may be optional and strategic.[5]

Viewed from this perspective, Experiment 1 substantially extends the results reported by Karlin and Kestenbaum (1968). The similarity between our results and theirs demonstrates that, contrary to what Pashler (1994a) argued, mean RTs from simple-reaction and choice-reaction tasks may be aptly compared, at least with respect to the effects of Task 2 response-selection difficulty on mean RTs in the PRP procedure. Indeed, it now seems even more likely that the underadditivity reported by Karlin and Kestenbaum (see Figure 2A) stemmed from concurrent response-selection processes, as some previous critics of the RSB hypothesis have inferred (Keele, 1973; Meyer & Kieras, 1997a, 1997b, 1999; Meyer et al., 1995).

Furthermore, the results of Experiment 1 disarm claims made by Van Selst and Jolicoeur (1997). In a series of studies with the PRP procedure, they attempted to replicate the results of Karlin and Kestenbaum (1968) but failed to do so. This led Van Selst and Jolicoeur to favor the RSB hypothesis and to discredit other theorists who had rejected it on the basis of Karlin and Kestenbaum's results. Nevertheless, our first experiment shows that Karlin and Kestenbaum's results are replicable and generalizable. Unlike Van Selst and Jolicoeur, we had no trouble obtaining underadditive SOA by S-R numerosity effects on mean Task 2 RTs for an auditory-manual secondary task paired with a visual-manual primary task.

Meyer and Kieras's (1997a, 1997b, 1999) AEC models (see Figure 3) provide some plausible explanations of the discrepancy between our results and those of Van Selst and Jolicoeur (1997). For example, in Van Selst and Jolicoeur's Experiment 2, their participants had mean Task 2 RTs that were more than 100 ms longer than those of Karlin and Kestenbaum's (1968) participants. These longer RTs, which
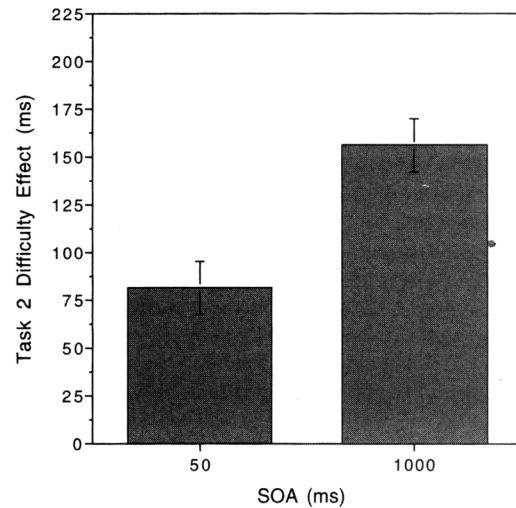


*Figure 5.* Response-selection difficulty effects and standard error bars for mean Task 2 reaction times at the shortest and longest stimulus onset asynchronies (SOAs) in Experiment 1. Standard error bars were derived from the mean squared error of the $F$ test for determining the reliability of the difference between the heights of the bars.

were produced when the secondary task was relatively difficult, occurred even after Task 1 had been finished. They imply that Van Selst and Jolicoeur's participants found their difficult Task 2 especially challenging and so perhaps became biased to adopt a cautious scheduling strategy (i.e., one with an early Task 2 lockout point). Such biases also may have prevailed because some of these participants—unlike those in our Experiment 1—were unusually familiar with the PRP procedure and had considerable a priori faith in the RSB hypothesis.

Attempting to alleviate the latter concern, Van Selst and Jolicoeur (1997, Experiment 3) conducted a subsequent study with naive, inexperienced participants and obtained results (i.e., additive SOA by S-R numerosity effects) like those from their previous studies. However, this subsequent study involved only one session, which—as we show later—may have precluded participants from learning to adopt a daring scheduling strategy. Thus, the main theoretical conclusions from our first experiment stand.

Yet another possible interpretation of the results from Experiment 1 must be considered too. Perhaps the interactions observed there involved mental operations other than response selection (viz., stimulus encoding). This is conceivable given that we manipulated response-selection difficulty initially by using a continuous sensory dimension (tone frequency) to vary the number of S-R pairs (i.e., tone-keypress associations) that Task 2 included. Participants

---

[5] The effects of Task 2 response-selection difficulty and SOA on Task 1 error rates complicate the interpretation of this interaction for Task 2 RTs. However, such complications did not occur in any of our other experiments and therefore cannot be used to discount the Task 2 RT data.

may have discriminated an 1120-Hz tone from a 1650-Hz tone more quickly when the tones were presented as one of two alternatives rather than as one of four. If so, then by locus-of-slack logic, the RSB hypothesis could explain the underadditive effects of S-R numerosity and SOA on mean Task 2 RTs because encoding processes for Tasks 1 and 2 might overlap temporally even though response-selection processes do not (see Figure 1).

It seems unlikely to us that the latter interpretation is empirically correct. Prior research (e.g., Brainard et al., 1962; Broadbent & Gregory, 1965; Morin & Forrin, 1965; Sternberg, 1969; Theios, 1973) has shown that S-R numerosity effects interact most strongly with those of other bona fide response-selection factors (e.g., S-R compatibility). By additive-factor logic (Sternberg, 1969), this suggests that S-R numerosity itself mainly affects response selection. Nevertheless, in our second experiment we explicitly dealt with this concern by changing the Task 2 stimuli from tones to printed digits. The high familiarity and discrete symbolic nature of digits further reduce the likelihood that manipulations of S-R numerosity involving them would affect stimulus encoding rather than response-selection difficulty. This enabled Experiment 2 to provide a more powerful test of our conclusions from Experiment 1.

## Experiment 2

The design of Experiment 2 was inspired by the previous research of Hawkins et al. (1979). For part of their research, they used a version of the PRP procedure in which Task 1 was an auditory-manual task with two S-R pairs and Task 2 was a visual-manual task with either two or eight S-R pairs. The S-R pairs of Task 1 involved tones and left-hand keypresses, whereas the S-R pairs of Task 2 involved printed digits and right-hand keypresses. Under these conditions, Hawkins et al. found a large underadditive interaction between the effects of SOA and S-R numerosity on Task 2 RTs. Their S-R numerosity effect averaged 180 ms at the longest SOA but only 35 ms at the shortest SOA (see Figure 2B). Assuming that these effects occurred during Task 2 response selection, the difference between them implies that response-selection processes for Tasks 1 and 2 overlapped temporally at the short SOAs. Of course, such overlap is consistent with the results of our Experiment 1 (see Figures 4 and 5). That Hawkins et al.'s S-R numerosity effects might have occurred instead through concurrent stimulus-encoding processes seems extremely unlikely, because the size of their interaction with SOA (i.e., 145 ms) far exceeds the known effects of stimulus numerosity on encoding processes for highly legible printed digits (e.g., Theios, 1973).

In Experiment 2, we followed Hawkins et al.'s (1979) PRP procedure closely, showing that their underadditive SOA and S-R numerosity effects could be replicated. Our approach further addressed two previously mentioned explanations that proponents of the RSB hypothesis (e.g., Pashler, 1994a) offered to discount the results of Karlin and Kestenbaum (1968). Because both Task 1 and Task 2 were always choice-reaction tasks throughout our second experiment, comparing results from simple and choice reactions—which

Karlin and Kestenbaum did—was not required here. Also, because the present stimuli for Task 2 were highly legible printed digits, S-R numerosity presumably had little, if any, effect on the duration of stimulus encoding. Instead, most of the numerosity effect found here is likely to have occurred during secondary-task response selection. By locus-of-slack logic, the underadditivity of this effect and that of SOA on mean Task 2 RTs therefore provides more evidence of concurrent response-selection processes, again supporting the AEC models (see Figure 3) of Meyer and Kieras (1997a, 1997b, 1999).

Another, related objective of Experiment 2 was to study the effects of practice on dual-task performance. To this end, in Experiment 2 we included three test sessions per participant, whereas each participant in Experiment 1 had served for only two sessions. The inclusion of a third session provided potential opportunities for observing practice-induced changes in task scheduling and for testing various tentative predictions based on the present AEC models. That such changes might occur here and yield relevant tests was suggested by results from some other research in our laboratory in which participants' preferred scheduling strategies depended on what types of practice they had (Meyer et al., 1995; Schumacher et al., 1997).

How might a third session of practice affect dual-task performance under the PRP procedure? One possibility was that *intratask automaticity* might be promoted (Brown & Carr, 1989), facilitating some stage(s) of processing within each task. For example, continued practice could help automatize the Task 2 response-selection stage, thereby decreasing its mean duration (Pashler & Baylis, 1991a). If so, then a Task 2 in which response selection is initially difficult (e.g., involves many S-R pairs) presumably would benefit more than would a Task 2 in which response selection is initially easy (e.g., involves few S-R pairs), because the difficult Task 2 offers more room for improvement. In turn, as practice progresses during a PRP experiment, this particular benefit would attenuate an underadditive interaction between the effects of SOA and response-selection difficulty on mean Task 2 RTs. Such attenuation would happen because decreases in the duration of Task 2 response selection tend to be manifested more at longer SOAs, where temporal processing slack cannot absorb them (Schweickert & Boggs, 1984).

Furthermore, it is possible that continued practice at dual-task performance might promote the acquisition of efficient *intertask combination strategies* (Brown & Carr, 1989), enhancing the coordination of processing stages across two or more tasks. For example, according to our AEC models (see Figure 3), task scheduling under the PRP procedure ordinarily could be affected by practice in one or another of three ways: (a) Participants might adopt a cautious (i.e., postponed Task 2 response selection) scheduling strategy initially and use it throughout practice; (b) participants might adopt a daring (i.e., unpostponed Task 2 response selection) scheduling strategy initially and use it throughout practice; or (c) participants might adopt a cautious scheduling strategy initially, but after some prac-

tice, switch to a daring scheduling strategy.[6] Given these first two possibilities and disregarding other types of practice effects, the size of the interaction between the SOA and response-selection difficulty effects on mean Task 2 RTs would stay the same as practice progresses, because the temporal overlap between response-selection processes for Tasks 1 and 2 would stay the same. However, the third possibility (i.e., a transition from cautious to daring scheduling) would yield an increasingly underadditive interaction, because the temporal overlap between response-selection processes would increase. As discussed elsewhere (Meyer & Kieras, 1997a, 1997b, 1999; Meyer et al., 1995), which of the preceding possibilities happens in practice may depend on various contextual factors, including the perceived difficulty of Tasks 1 and 2, the extent to which they involve the same perceptual or motor modalities, the details of instructions about relative task priorities, and the personality traits of individual participants. The nature of these dependencies is considered in more detail later.

Meanwhile, Table 2 summarizes net results that the aforementioned potential changes in intratask automaticity and intertask combination strategies would have jointly during dual-task practice under the PRP procedure. Here the rows of the table indicate whether the response-selection difficulty effect in Task 2 remains constant or decreases with automatization. The columns of the table indicate whether task scheduling is consistently cautious, initially cautious but subsequently daring because of increasing strategic efficiency, or consistently daring. Entries in the table's cells indicate what the conjunctions of these alternative possibilities yield in terms of an interaction between the effects of SOA and response-selection difficulty on mean Task 2 RTs. As they show, the interaction may become less underadditive, remain constant, or become more underadditive through practice, depending on the relative contributions of evolving intratask automaticity and intertask combination strategies.

With Experiment 2, we confirm that some representative circumstances indeed yield this first possible pattern (i.e., decreasing underadditive interaction). Subsequently, in Experiment 3 we provide evidence that by principled adjust-

ments to prevailing task conditions, the opposed third pattern (i.e., increasing underadditive interaction) also can occur. That the obtained interaction may be further manipulated in accord with our AEC models is demonstrated in Experiment 4. Taken together, the results from the following experiments strongly support these models, which constitute a powerful and comprehensive theoretical framework for understanding and predicting a multiplicity of practice effects on dual-task performance. In contrast, such effects further discredit the traditional RSB hypothesis, which implies that SOA and response-selection difficulty in Task 2 should always affect mean Task 2 RTs additively regardless of practice, because the putative RSB is claimed to be structural and immutable rather than strategic and flexible (Pashler, 1994a).

## Method

### Participants

Ten right-handed undergraduate students (8 men and 2 women) at the University of Michigan participated as paid volunteers. They came from the same population as those in Experiment 1 but had not been tested previously. All participants had normal or corrected-to-normal vision and were paid $4.50 per hour plus a bonus based on the quality of their performance.

### Apparatus

The apparatus was the same as that used in Experiment 1.

### Design and Procedure

*Tasks.* Participants performed two tasks. Task 1 was an auditory-manual task. On each Task 1 trial, participants heard either an 800-Hz or 1200-Hz tone and responded by pressing either the left index-finger or left middle-finger key, respectively. Task 2 was a visual-manual task. On each Task 2 trial, one of eight digits (2 through 9) appeared in the center of the display monitor. If a 2, 5, 6, or 9 appeared, participants pressed the right index-finger key. If a 3, 4, 7, or 8 appeared, participants pressed the right middle-finger key. Task 2 had two (easy and hard) levels of response-selection difficulty. The easy version of Task 2 involved two S-R pairs that used the digits 2 and 3, whereas the hard version involved eight S-R pairs that used the digits 2 through 9. The stimuli for the two tasks were separated by one of five SOAs: 50, 150, 250, 500, or 1,000 ms.

*Sessions.* Three test sessions were conducted on separate days for each participant. Session 1 included 13 trial blocks. The first three of these were single-task blocks, with one block of 24 trials for Task 1 and one block of 24 trials for each level of Task 2 difficulty. The next three blocks were dual-task blocks involving Task 1 and the easy Task 2; successively across these blocks, one of them contained 20 trials with a 1,000-ms SOA, one contained 20

Table 2
*Expected Alternative Practice-Induced Changes in Interaction of SOA and Task 2 Response-Selection Difficulty Effects on Mean Task 2 RTs Under the AEC Models*

| Difficulty effect | Scheduling strategies used early/late in practice | | |
|---|---|---|---|
| | C/C | C/D | D/D |
| Constant | 0 | ↑ | X |
| Decreasing | 0 | ↑ | ↓ |

*Note.* Entries in cells of table denote underadditive interaction with increasing absolute value (↑), constant underadditive interaction (X), underadditive interaction with decreasing absolute value (↓), or constant null interaction (0). C = conservative task scheduling; D = daring task scheduling; SOA = stimulus onset asynchrony; RT = reaction time; AEC = adaptive executive control.

[6] In the present context, a fourth conceivable scenario, namely, a transition from daring to cautious scheduling, is antithetical to heuristic task-scheduling principles associated with our AEC models (Meyer & Kieras, 1997a, 1997b, 1999; Meyer et al., 1995). Once participants have adopted a daring scheduling strategy during the course of practice here, there ordinarily would be no utilitarian reason for them later to adopt cautious scheduling.

trials with a 150-ms SOA, and one contained 40 trials with a 500-ms SOA. The remaining dual-task blocks of the session each contained 80 trials and involved all five SOAs. Across these blocks, the level of Task 2 difficulty alternated.

Sessions 2 and 3 each included 13 dual-task blocks with 80 trials per block. The first block in each of these sessions involved Task 1 and the hard Task 2 with a 500-ms SOA on each trial. The remaining 12 blocks of each session used all five SOAs. During these blocks, the easy or hard version of Task 2 occurred in successive groups of three blocks each. For all participants, the first group of blocks with the easy Task 2 preceded the first group of blocks with the hard Task 2, and then the groups repeated alternately. All of Session 1 and the first blocks of Sessions 2 and 3 were for preliminary warm-up and were omitted from subsequent data analyses.

*Trial blocks.* All relevant tone–digit–SOA combinations occurred equally often within each trial block. Participants were told before each block which task(s) would be involved.

*Trials.* The trials of each block began with a 100-Hz warning tone for 100 ms accompanied by a fixation cross in the center of the display monitor. On dual-task trials, 500 ms after the offset of the warning tone, a Task 1 tone sounded for 40 ms, and after the SOA, a Task 2 digit replaced the fixation cross. On single-task trials, only one stimulus was presented. The assignments of stimuli and SOAs to particular trials were randomized within each block.

*Feedback and points.* After each trial, participants received point totals that depended on their performance. For Task 1, we awarded 200 points per correct response and deducted 2 points per every 10 ms taken to respond correctly; 200 points were deducted for every incorrect Task 1 response. For Task 2, we awarded 100 points per correct response and deducted 1 point per every 10 ms taken to respond correctly; 100 points were deducted for every incorrect Task 1 response. On dual-task blocks, participants had to make their Task 1 response first or else both responses were considered to be incorrect.[7] Participants earned $1 for every 20,000 points they scored. They were fully informed about the reward system before the experiment began.

After each trial block, participants received detailed feedback about their numbers of correct responses, mean RTs, and points. Also, for the first five blocks of Session 1 and the first block of Sessions 2 and 3, participants received feedback about response accuracy and points after each trial. Subsequently, feedback about response accuracy occurred only after trials for which there were incorrect responses.

## Results

Outliers were trimmed and remaining RT data were analyzed in the same way as those from Experiment 1, with session included as an additional factor. The outlier-trimming procedure removed 4.7% of the original data set, leaving 17,212 response pairs. Figure 6 shows the mean RTs of correct responses for each task in Sessions 2 and 3 as a function of Task 2 response-selection difficulty and SOA. Table 3 shows the corresponding error rates. In what follows, we discuss the results more fully for each task.

### Task 1

*Reaction times.* The main effect of session on mean Task 1 RTs was reliable, $F(1, 9) = 19.82, p < .005$. Across sessions, mean Task 1 RTs decreased from 360 to 330 ms. In contrast, neither the main effect of Task 2 response-selection

difficulty, $F(1, 9) = 0.02, p > .85$, nor that of SOA, $F(4, 36) = 2.00, p > .11$, was reliable. However, a small but reliable interaction occurred between the effects of session and Task 2 difficulty, $F(1, 9) = 5.55, p < .05$. Mean RTs for Task 1 when paired with the easy Task 2 were 9 ms slower in Session 2 and 7 ms faster in Session 3 than when Task 1 was paired with the hard Task 2. There were no other reliable effects ($p > .40$ in all cases).

*Error rates.* The overall Task 1 error rate was low (1.97%). The only reliable contrast was an interaction between the effects of session and SOA, $F(4, 36) = 3.60, p < .05$. Error rates increased as SOA decreased in Session 2, but error rates did not change reliably with SOA in Session 3.

### Task 2

*Reaction times.* All main effects and interactions for mean Task 2 RTs were reliable. Across sessions, mean Task 2 RTs decreased from 503 to 475 ms, $F(1, 9) = 15.22, p < .005$. Mean RTs were faster when Task 2 was easy (436 ms) than when Task 2 was hard (543 ms), $F(1, 9) = 148.09, p < .0005$. The effect of Task 2 response-selection difficulty decreased from 116 ms in Session 2 to 97 ms in Session 3, $F(1, 9) = 9.26, p < .05$. There was a moderate SOA effect, $F(4, 36) = 94.59, p < .0005$ (see Figure 6). The SOA effect decreased from Session 2 to Session 3, $F(4, 36) = 8.26, p < .0005$. Similarly, a reliable interaction occurred between the effects of Task 2 response-selection difficulty and SOA, $F(4, 36) = 17.35, p < .0005$; the difficulty effect decreased as SOA decreased. Finally, the three-way interaction between the effects of session, Task 2 difficulty, and SOA was reliable, $F(4, 36) = 2.97, p < .05$. This stemmed from a reliable decrease in the Task 2 difficulty effect at the longest SOA as practice progressed, consistent with the entry in the lower right cell of Table 2 and with the present AEC models.

Figure 7 shows the response-selection difficulty effects on mean Task 2 RTs at the shortest and longest SOAs for Sessions 2 and 3. The differences between the difficulty effects at these SOAs (64 ms in Session 2 and 44 ms in Session 3) were reliable during each session: $t(9) = 5.69, p < .0005$ for Session 2, and $t(9) = 5.97, p < .0005$ for Session 3.

*Error rates.* The overall Task 2 error rate was 4.03%. There was a reliable main effect of Task 2 response-selection difficulty, $F(1, 9) = 17.16, p < .005$. Participants made more errors when Task 2 was hard (5.23%) than when Task 2 was easy (2.82%). The effect of SOA was also reliable, $F(4, 36) = 7.86, p < .0005$; error rates increased as SOA decreased. No other effects on Task 2 error rates were reliable.

### Discussion

In Experiment 2 we replicated the results of both Experiment 1 and Hawkins et al. (1979). The effect of response-selection difficulty on mean Task 2 RTs was 64 ms less at the shortest SOA than at the longest SOA in Session 2, and 44

---

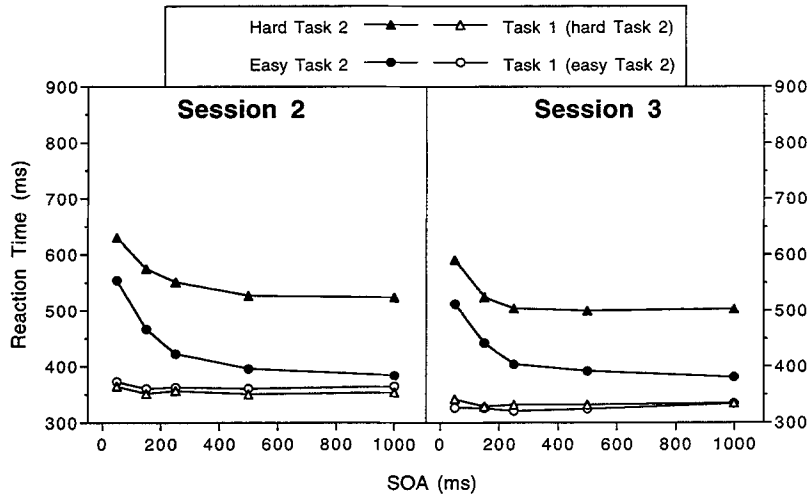[7] Out-of-order errors occurred on fewer than 0.09% of trials overall.

*Figure 6.* Mean reaction times for Tasks 1 and 2 as a function of Task 2 response-selection difficulty and stimulus onset asynchrony (SOA) in Sessions 2 and 3 of Experiment 2.

ms less in Session 3 (see Figure 7). These underadditive interactions are difficult to explain with the traditional RSB hypothesis. Instead, they strongly suggest that response-selection processes for Tasks 1 and 2 temporally overlapped at the shorter SOAs. This again supports the present AEC models, showing that insofar as RSBs exist, they are probably strategic and flexible rather than structural and immutable.

More specifically, the repeated occurrence of underadditive interactions during Sessions 2 and 3 implies that participants of Experiment 2 adopted a daring scheduling strategy early in practice and continued to apply it consistently thereafter. According to heuristic principles of task scheduling associated with AEC models (Meyer & Kieras, 1997a, 1997b, 1999; Meyer et al., 1995), such consistency

Table 3
*Mean Error Rates (%) in Sessions 2 and 3 of Experiment 2 for Each Task as a Function of Task 2 Response-Selection Difficulty and Stimulus Onset Asynchrony (SOA)*

| Session and Task 2 type | SOA (ms) | | | | |
|---|---|---|---|---|---|
| | 50 | 150 | 250 | 500 | 1,000 |
| | Task 1 | | | | |
| Session 2 | | | | | |
| Hard | 2.50 | 1.88 | 1.56 | 1.35 | 1.35 |
| Easy | 3.44 | 2.60 | 1.67 | 2.40 | 1.25 |
| Session 3 | | | | | |
| Hard | 1.67 | 1.46 | 1.98 | 1.67 | 1.98 |
| Easy | 2.50 | 2.29 | 2.50 | 0.94 | 2.50 |
| | Task 2 | | | | |
| Session 2 | | | | | |
| Hard | 6.56 | 4.79 | 6.04 | 3.96 | 4.90 |
| Easy | 3.96 | 4.17 | 3.44 | 2.08 | 2.40 |
| Session 3 | | | | | |
| Hard | 6.46 | 5.94 | 5.00 | 3.65 | 5.00 |
| Easy | 3.02 | 2.92 | 3.33 | 1.56 | 1.35 |

may have resulted because mean Task 1 RTs were shorter than those for both the easy and hard versions of Task 2 (e.g., 360 vs. 384 and 524 ms, respectively, at the longest SOA in Session 2). With relatively fast Task 1 responses being possible, participants did not especially need to concern themselves about whether Task 2 responses would precede Task 1 responses at the shortest SOA and thus violate instructions of the PRP procedure. This could have encouraged the use of an executive process that let response selection be completed for Task 2 before progress on Task 2 was postponed temporarily.

Interestingly, Experiment 2 also yielded a reliable triple interaction between the effects of SOA, response-selection difficulty, and session on mean Task 2 RTs. In going from Session 2 to Session 3, the difficulty effect decreased more at the longest SOA than at the shortest SOA. As mentioned already (see Table 2, lower right cell), this pattern is what should happen when intratask automatization (Brown & Carr, 1989) causes the difficulty effect to decrease under conditions in which participants adopt and apply a daring scheduling strategy consistently throughout practice. Further supporting our AEC models, these results show that response-selection processes for Tasks 1 and 2 do not have to be fully automatized before they can overlap temporally. Rather, such overlap—and the daring task scheduling that enables it—may ensue early in practice as a characteristic of dual-task performance distinct from intra-task automatization.

Of course, ardent advocates of immutable structural RSBs still might argue that Experiment 2's results stemmed from effects of S-R numerosity on the duration of stimulus encoding for Task 2 (Pashler, 1984, 1994a; Pashler & Baylis, 1991b). If so, then the traditional RSB hypothesis perhaps could be maintained. However, such arguments lose much of their force after careful inspection of findings from many previous studies that have used discrete symbolic stimuli
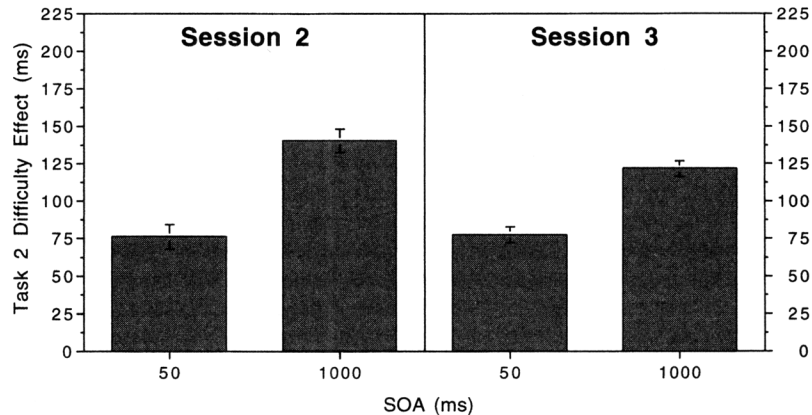
*Figure 7.* Response-selection difficulty effects and standard error bars for mean Task 2 reaction times at the shortest and longest stimulus onset asynchronies (SOAs) in Sessions 2 and 3 of Experiment 2. Standard error bars were derived from the mean squared error of the *F* test for determining the reliability of the difference between the heights of the bars.

and manipulated S-R numerosity to influence response-selection difficulty.

For example, consider a study by Sternberg (1969, Experiment 5). He orthogonally manipulated S-R numerosity, S-R compatibility, and visual stimulus discriminability (intact vs. degraded presentation) in a digit-naming task. His participants produced an interaction of about 90 ms between the effects of S-R numerosity and S-R compatibility on mean RTs, which was almost 27% of the overall RT magnitude. Because the S-R compatibility effect presumably occurred during response selection (Kornblum, Hasbroucq, & Osman, 1990; McCann & Johnston, 1992; Sanders, 1980; Sternberg, 1969), this large interaction implies that much of the S-R numerosity effect likewise took place there.

In contrast, Sternberg's participants produced only about a 24-ms interaction between the effects of S-R numerosity and stimulus discriminability on mean RTs, which was just 7% of the overall RT magnitude. Because the discriminability effect presumably occurred during stimulus encoding (Sanders, 1980; Sternberg, 1969), this small interaction implies that little of the S-R numerosity effect took place there. As Sternberg (1969, p. 301) concluded, "One might want to argue from the relative weakness of the interaction [between S-R numerosity and stimulus discriminability] to the relative weakness of the effect of [S-R numerosity] on the [encoding] stage." Also supporting this conclusion, other researchers have found only tiny or null effects of S-R numerosity on the duration of stimulus encoding (Brainard et al., 1962; Morin & Forrin, 1965; Theios, 1973).[8]

Furthermore, even Sternberg's (1969, Experiment 5) 24-ms interaction between the effects of S-R numerosity and stimulus discriminability may have been an artifact of the experimental design from which they came. The design condition with two S-R pairs required participants to distinguish between the digits 1 and 8, which have distinctive perceptual features (lines vs. curves) that could facilitate stimulus encoding in this condition. These features were less

distinctive in the condition with eight S-R pairs. As a result, this could have caused the manipulation of S-R numerosity to affect stimulus encoding indirectly. Better control of perceptual-feature distinctiveness across the levels of S-R numerosity might have yielded an even smaller or null interaction between S-R numerosity and discriminability factors in Sternberg's study.

Given the preceding considerations, two strong inferences follow from the results of Experiment 2: (a) Our manipulation of S-R numerosity in the secondary task had much, if not all, of its effect on the duration of response selection for Task 2, and (b) the large interaction of this effect and that of SOA on mean Task 2 RTs occurred because responses for Tasks 1 and 2 were selected concurrently at short SOAs.[9]

To support these conclusions further, in our third experiment we introduced another complementary factor, S-R compatibility, which presumably influences response selection rather than stimulus encoding or movement production. Again we showed that the manipulation of such a factor in Task 2 sometimes yields effects that interact underadditively

[8] More recently, Sternberg (1998) interpreted these and subsequent results from his laboratory to suggest a larger, more robust S-R numerosity effect on stimulus encoding. However, based on our inspection of many other relevant studies in the literature, we continue to maintain that this effect is typically small, especially relative to the usual effects of S-R numerosity on response selection.

[9] Of course, these conclusions rest on our prior assumption (see Footnote 1) that stimulus-encoding processes in Experiment 2 were similar to those used by participants in previously cited studies that manipulated S-R numerosity (viz., Brainard et al., 1962; Morin & Forrin, 1965; Sternberg, 1969; Theios, 1973). That stimulus-encoding processes might be similar in these studies and ours is highly plausible. As stated previously, the present and past S-R numerosity manipulations differ mostly in the response characteristics of the tasks, which affect late processing stages (e.g., movement production; Miller & Ulrich, 1998) and presumably have little effect on an early stimulus-encoding stage.

with those of SOA on mean Task 2 RTs, manifesting concurrent response-selection processes. However, unlike before, Experiment 3 also revealed that this underadditivity may emerge after an initial practice period in which participants' strategy of task scheduling is cautious rather than daring. This new discovery reinforced our previous inferences about the mutability of RSBs and the veracity of AEC models.

## Experiment 3

In Experiment 3, Task 1 was an auditory-manual task similar to the one used in Experiment 2. Task 2 was a visual-manual task with keypress responses designated by spatial locations of symbolic stimuli. We manipulated response-selection difficulty for Task 2 by varying the compatibility of the S-R pairs there. As part of this manipulation, mappings from stimulus locations to keypress responses were either spatially ordered or haphazard. Many prior studies have varied S-R compatibility in this way in order to influence response-selection difficulty (Duncan, 1977; Fagot & Pashler, 1992; Fitts & Seeger, 1953; Kornblum et al., 1990; McCann & Johnston, 1992).

Following our previous experiments, the rationale of Experiment 3 is straightforward. If the RSB hypothesis is correct, then throughout the sessions of this experiment, the S-R compatibility effect on mean Task 2 RTs should be additive with the effect of SOA, regardless of how much practice participants have received. This is because all of the compatibility effect presumably occurs in a stage of processing (i.e., response selection for Task 2) that always takes place after the temporal slack caused by the putative immutable bottleneck (McCann & Johnston, 1992; Pashler, 1994a). However, if response-selection processes for Tasks 1 and 2 can overlap temporally, then contrary to the RSB hypothesis but consistent with our AEC models, an underadditive interaction should emerge again at some point during practice in Experiment 3. Unlike in Experiments 1 and 2, potential confoundings of factor effects on stimulus encoding and response selection are avoided here, given that stimulus encoding does not appear to be influenced whatsoever by S-R compatibility (Sanders, 1980; Sternberg, 1969). Thus, by seeking underadditive effects of SOA and S-R compatibility—which determines response-selection difficulty—in Experiment 3, we may test competing theoretical alternatives further with a powerful complementary manipulation.

Given heuristic principles of task scheduling, this manipulation also might likely be expected to yield a different pattern of practice effects than was found during Experiment 2. Early in practice, the inherent subjective difficulty of the incompatible Task 2 S-R mapping could lead participants to adopt a cautious (postponed Task 2 response selection) scheduling strategy, because they perhaps would surmise that it is difficult to perform Task 2 at the same time as Task 1. Nevertheless, after more experience with the nuances of both tasks, they could realize that such concurrent performance is possible and so shift to a daring (unpostponed Task 2 response selection) scheduling strategy. If so, then as

practice progresses during Experiment 3, the effects of SOA and S-R compatibility (i.e., response-selection difficulty) on mean Task 2 RTs should become more underadditive, whereas they became less underadditive during Experiment 2 when participants adopted a daring scheduling strategy at the outset. Such increasing underadditivity may occur despite intratask automatization's causing the main effect of response-selection difficulty to decrease as practice progresses (see Table 2, lower middle cell). Consequently, when combined with Experiment 2, our third experiment has the potential for demonstrating separable independent contributions of intratask automatization and intertask combination strategies to dual-task performance (cf. Brown & Carr, 1989), consistent with implications of the present AEC models.

## Method

### Participants

Eight right-handed undergraduate students (4 men and 4 women) from the University of Michigan participated in this study as paid volunteers. They came from the same population as those in the previous experiments but had not been tested previously. Participants had normal or corrected-to-normal vision and were paid $5.00 per hour plus a bonus based on the quality of their performance.

### Apparatus

The apparatus was the same as that used in previous experiments.

### Design and Procedure

*Tasks.* Participants performed two tasks. Task 1 was an auditory-manual task. On each Task 1 trial, participants heard either an 1120-Hz or a 1450-Hz tone and responded by pressing the left middle-finger or left index-finger key, respectively. Task 2 was a visual-manual task. On each Task 2 trial, an O replaced one of four dashes in a horizontal row centered on the display monitor. Task 2 had two (easy and hard) levels of response-selection difficulty. In the easy Task 2, participants pressed the right index, middle, ring, or little finger keys when the O appeared in the far left, middle left, middle right, or far right spatial positions, respectively. In the hard Task 2, participants pressed the right index, middle, ring, or little finger keys when the O appeared in the middle left, far right, far left, or middle right positions, respectively. The stimuli for the two tasks were separated by one of five SOAs: 50, 150, 250, 500, or 1,000 ms.

*Sessions.* Three test sessions were conducted on separate days for each participant. Session 1 included 14 single-task blocks and 8 dual-task blocks with the auditory-manual task and various versions of the visual-manual task.

Sessions 2 and 3 each included 8 single-task blocks and 16 dual-task blocks. The dual-task blocks were paired such that for each block pair, either the easy or hard Task 2 had to be performed, and the difficulty of Task 2 alternated across block pairs in a fashion that was counterbalanced between participants. Preceding each block pair was a single-task block involving the version of Task 2 to be performed next. Throughout the experiment, there were 20 trials per single-task block and 40 trials per dual-task block. All of Session 1 and the single-task blocks of Sessions 2 and 3 were for

preliminary warm-up and were omitted from subsequent data analyses.

*Trial blocks.* All relevant tone–location–SOA combinations occurred equally often within each trial block. Participants were told before each block which task(s) would be involved.

*Trials.* Each trial began with an 880-Hz warning tone for 50 ms and a row of four dashes presented in the center of the display monitor. On dual-task trials, 500 ms after the offset of the warning tone, a Task 1 tone sounded for 40 ms, and after the SOA, a Task 2 stimulus replaced one of the dashes. On single-task trials, only one task stimulus was presented. The assignments of stimuli and SOAs to particular trials were randomized within each block.

*Feedback and points.* The protocols for providing feedback and points were identical to those in Experiment 1 except that participants earned $1 for every 40,000 points they scored. After each trial block, participants received detailed feedback about their numbers of correct responses, mean RTs, and points. During the first 10 blocks of Session 1, participants received feedback about response accuracy and points after each trial. Subsequently, participants received this feedback only after trials on which there were incorrect responses.[10]

## Results

Outliers were trimmed and remaining data were analyzed in the same way as those in Experiment 2. The outlier-trimming procedure removed 5.2% of the original data set, leaving 9,284 response pairs. Figure 8 shows the mean RTs of correct responses for each task in Sessions 2 and 3 as a function of Task 2 response-selection difficulty (i.e., S-R compatibility) and SOA. Table 4 shows the corresponding error rates. In what follows, we discuss the results more fully for each task.

### Task 1

*Reaction times.* The main effect of session on Task 1 RTs was reliable, $F(1, 7) = 12.24, p < .05$. Across sessions, mean Task 1 RTs decreased from 430 to 401 ms. Neither the main effect of Task 2 response-selection difficulty, $F(1, 7) = 1.95, p > .20$, nor that of SOA, $F(4, 28) = 2.02, p > .10$, was reliable. There was a reliable interaction between the effects of session and Task 2 difficulty, $F(1, 7) = 5.39, p < .05$. The Task 2 difficulty effect on mean Task 1 RTs decreased from 20 ms in Session 2 to 5 ms in Session 3. There was also a small but reliable interaction between the effects of Task 2 difficulty and SOA, $F(4, 28) = 2.83, p < .05$. The effect of Task 2 difficulty on mean Task 1 RTs increased as SOA decreased. There were no other reliable interactions.

*Error rates.* The overall Task 1 error rate was fairly low (1.66%). The only reliable interaction was between the effects of session and Task 2 response-selection difficulty, $F(1, 7) = 7.09, p < .05$. Participants made more Task 1 errors in Session 2 and fewer in Session 3 when performing the hard Task 2 than when performing the easy Task 2, but the error rates were less than 3% in all cases.

### Task 2

*Reaction times.* All main effects on Task 2 RTs were reliable. Across sessions, mean Task 2 RTs decreased from

612 to 567 ms, $F(1, 7) = 7.29, p < .05$; Task 2 responses were faster when Task 2 was easy (514 ms) than when Task 2 was hard (664 ms), $F(1, 7) = 56.72, p < .0005$; and there was a large SOA effect, $F(4, 28) = 43.05, p < .0005$ (see Figure 8). The interaction between the effects of session and Task 2 response-selection difficulty was reliable, $F(1, 7) = 16.26, p < .01$. Across sessions, the Task 2 difficulty effect decreased from 175 ms to 125 ms. The interaction between the effects of Task 2 difficulty and SOA was not reliable, $F(4, 28) = 2.02, p > .10$, but a reliable three-way interaction occurred among the effects of session, Task 2 difficulty, and SOA, $F(4, 28) = 2.74, p < .05$.

We also analyzed mean Task 2 RTs separately for each session. All main effects on them were reliable for each session ($p < .0005$ in all cases). The interaction between the effects of Task 2 response-selection difficulty and SOA was not reliable for Session 2, $F(4, 28) = 0.62, p > .65$, but was reliable for Session 3, $F(4, 28) = 4.42, p < .01$. Figure 9 shows the response-selection difficulty effects on mean Task 2 RTs at the shortest and longest SOAs for Sessions 2 and 3. The 51-ms difference between the difficulty effects at these SOAs in Session 3 was reliable, $t(7) = 3.72, p < .01$.

*Error rates.* The overall Task 2 error rate was 2.54%. Only SOA affected it reliably, $F(4, 28) = 3.10, p < .05$. Participants made more errors at the 50-ms SOA than at any other SOA.

## Discussion

In Experiment 3, we obtained results with implications that confirm and extend the results of our previous experiments. During Session 3 of Experiment 3, SOA and S-R compatibility for Task 2 had substantially underadditive effects on mean Task 2 RTs (see Figures 8 and 9). The size of this underadditivity (51 ms) was similar to the underadditivity that occurred on average during Experiment 1 (79 ms) and Experiment 2 (54 ms), where we manipulated S-R numerosity for Task 2. Earlier, concerns were raised about whether these prior underadditivities stemmed from concurrent response-selection processes, because the functional loci of S-R numerosity effects are controversial (cf. Brainard et al., 1962; Morin & Forrin, 1965; Sanders, 1980; Sternberg, 1969, 1998; Theios, 1973). Resolving such concerns, S-R compatibility is a prototypical factor that selectively influences response-selection processes (Fagot & Pashler, 1992; Fitts & Seeger, 1953; Kornblum et al., 1990; McCann & Johnston, 1992; Sanders, 1980; Sternberg, 1969). The similarity of its present underadditive effects to those of S-R numerosity suggests that both of these factors influence response-selection processes and that these processes may overlap temporally at short SOAs for Tasks 1 and 2 of the

---

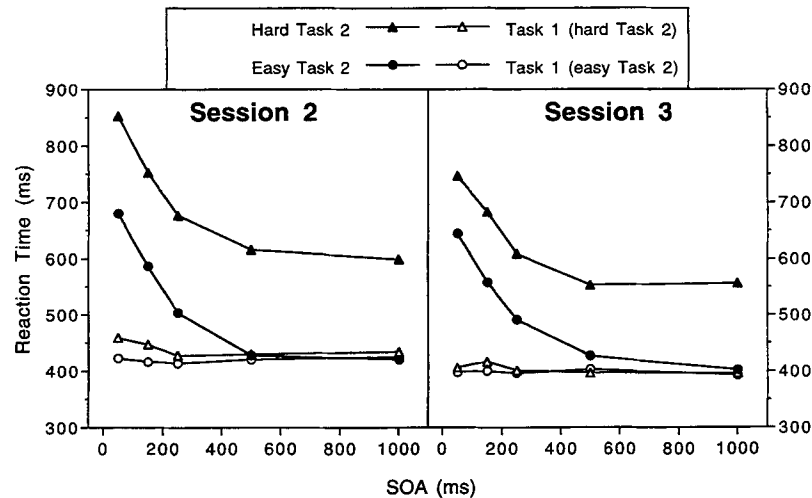[10] Out-of-order errors occurred on fewer than 0.06% of trials overall.

*Figure 8.* Mean reaction times for Tasks 1 and 2 as a function of Task 2 response-selection difficulty and stimulus onset asynchrony (SOA) in Sessions 2 and 3 of Experiment 3.

PRP procedure. This outcome casts more doubt on the RSB hypothesis while further supporting our AEC models.

### Theoretical Significance of Practice Effects

Our AEC models are supported as well by other comparisons between the results of Experiments 2 and 3. On the basis of heuristic task-scheduling principles (Meyer & Kieras, 1997a, 1997b, 1999; Meyer et al., 1995), we hypothesized that Experiment 3 might yield a different pattern of practice effects than Experiment 2 did. According to the latter hypothesis, participants in Experiment 3 were expected to adopt an initial cautious scheduling strategy (postponed Task 2 response selection) and practice with it before shifting to a daring scheduling strategy (unpostponed

Table 4

*Mean Error Rates (%) in Sessions 2 and 3 of Experiment 3 for Each Task as a Function of Task 2 Response-Selection Difficulty and Stimulus Onset Asynchrony (SOA)*

| Session and Task 2 type | SOA (ms) | | | | |
|---|---|---|---|---|---|
| | 50 | 150 | 250 | 500 | 1,000 |
| | Task 1 | | | | |
| Session 2 | | | | | |
| Hard | 3.13 | 1.37 | 1.37 | 1.95 | 2.15 |
| Easy | 2.54 | 1.56 | 1.37 | 0.98 | 0.78 |
| Session 3 | | | | | |
| Hard | 1.37 | 1.56 | 0.98 | 1.17 | 1.76 |
| Easy | 0.98 | 2.34 | 2.15 | 1.56 | 2.15 |
| | Task 2 | | | | |
| Session 2 | | | | | |
| Hard | 5.27 | 2.15 | 1.37 | 4.30 | 2.34 |
| Easy | 2.93 | 1.37 | 1.17 | 1.17 | 2.34 |
| Session 3 | | | | | |
| Hard | 2.73 | 3.13 | 2.34 | 2.54 | 4.88 |
| Easy | 2.54 | 1.95 | 2.15 | 1.76 | 2.34 |

Task 2 response selection), whereas participants in Experiment 2 adopted a daring scheduling strategy without much prior practice. This hypothesis seemed plausible because in Experiment 3, one version of Task 2 involved a highly incompatible S-R mapping, which at first could bias participants to prefer postponing response selection for Task 2 until Task 1 had been completed. The present results confirm these expectations. As practice progressed during Experiment 3, the joint effects of SOA and S-R compatibility on mean Task 2 RTs changed from a null to an underadditive interaction, consistent with the anticipated shift of scheduling strategy (see Table 2, middle column). The flexibility of task scheduling manifested by the different observed patterns of practice effects, which occurred along with persistent large PRP effects across experiments, is easily understood and explained in terms of our AEC models but not in terms of the traditional RSB hypothesis.

Of course, one might ask whether the increasing underadditivity of SOA and S-R compatibility effects in Experiment 3 simply stemmed from increasing intratask automaticity (cf. Brown & Carr, 1989). We believe the answer is "no." As we show subsequently in Experiment 4, even if Task 2 involves a highly incompatible S-R mapping, participants may adopt a daring scheduling strategy early in practice when other prevailing task conditions encourage this choice. Again consistent with our AEC models and with heuristic principles of task scheduling, the choice of a particular scheduling strategy is separable from other aspects of dual-task performance. Before pursuing these matters more fully, however, we should consider additional important details about the results of Experiment 3.

### Test for Multiple Immutable Structural Bottlenecks

In particular, comparing mean RTs from Sessions 2 and 3 of Experiment 3 lets us test a multiple structural-bottleneck hypothesis of De Jong (1993). He stated that dual-task
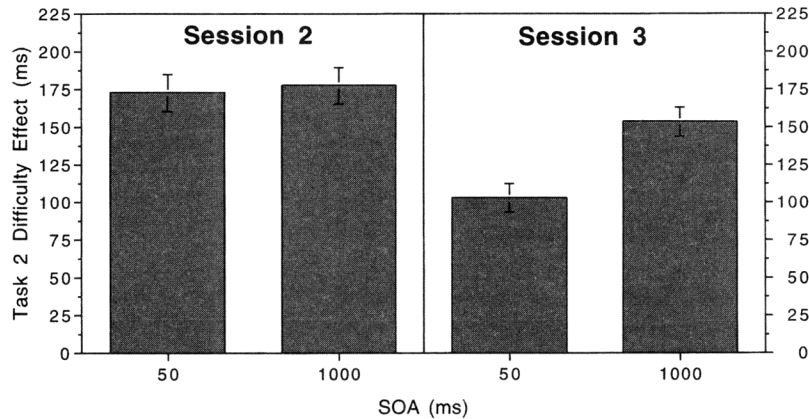
*Figure 9.* Response-selection difficulty effects and standard error bars for mean Task 2 reaction times at the shortest and longest stimulus onset asynchronies (SOAs) in Sessions 2 and 3 of Experiment 3. Standard error bars were derived from the mean squared error of the *F* test for determining the reliability of the difference between the heights of the bars.

performance is mediated by two types of immutable structural bottlenecks: one for response selection and another for the production of movements that involve the same (e.g., manual) response modality. This claim, which integrates the proposals of several past theorists (e.g., Keele, 1973; Pashler, 1984; Welford, 1980), was made to explain underadditive interactions between the effects of SOA and Task 2 response-selection difficulty on mean Task 2 RTs while preserving the essence of the traditional RSB hypothesis. As De Jong noted, some such interactions may be attributable to the limited capacity of a movement-production stage that follows a structural response-selection bottleneck and that converts abstract identities of selected responses (e.g., "press left index-finger key" and "press right middle-finger key") into motor features and muscle commands for executing overt response movements (cf. Keele, 1968; Meyer & Kieras, 1997a, 1997b; Rosenbaum, 1980).

According to the multiple structural-bottleneck hypothesis (De Jong, 1993), if both tasks of the PRP procedure involve the same response modality, then producing movements for Task 1 imposes a motor refractory period (MRP) on the production of movements for Task 2. By definition, the Task 1 MRP is the minimum possible time, which De Jong estimated to be about 200 ms, between the start of Task 1 movement production and the start of Task 2 movement production. Because of the Task 1 MRP, temporal slack would occur in progress on Task 2 whenever response selection for Task 2 finished before movement production for Task 1 was completed. This slack could absorb effects of response-selection difficulty on mean Task 2 RTs even though response selection for Task 2 also is constrained by a bottleneck and begins only after the Task 1 response has been selected. As a result, underadditive effects of SOA and response-selection difficulty might occur in the presence of an RSB, making them irrelevant for rejecting the RSB hypothesis.

Such irrelevant underadditivity can happen whenever (a)

Tasks 1 and 2 of the PRP procedure involve the same response modality, (b) there are relatively short as well as long SOAs, and (c) response selection for the easy version of Task 2 takes less time than the Task 1 MRP.[11] For example, consider the experiment by Karlin and Kestenbaum (1968). In that experiment, both Task 1 and Task 2 required manual responses, there were 50-ms and 1,150-ms SOAs, and the easy (simple reaction) Task 2 yielded a mean RT less than 200 ms at the longest SOA. Thus, these experimental conditions satisfied a priori assumptions of the multiple structural-bottleneck hypothesis, so the underadditive interaction that occurred there (see Figure 2A) does not definitively disprove the existence of an immutable RSB (De Jong, 1993).

The multiple structural-bottleneck hypothesis might explain the results from our Experiments 1 and 2 as well. In each of them, Tasks 1 and 2 required manual responses, there were short and long SOAs, and response selection for the easy Task 2 probably took less time than the putative Task 1 MRP. Together, these circumstances, like those of Karlin and Kestenbaum (1968), satisfy De Jong's (1993) assumptions about how underadditive SOA and response-selection difficulty effects could emerge in the context of an immutable RSB.

However, the results from our third experiment cannot be explained like this. During Session 3 of Experiment 3, the mean RT for the easy Task 2 at the longest SOA was virtually the same as during Session 2. This suggests that the nature and duration of response selection for the easy Task 2 did not

_____

[11] More precisely, for the multiple structural-bottleneck hypothesis to be applicable, the following inequality must hold: $T_{s2} < T_{m1} - \max [0, \text{SOA} + T_{e2} - T_{e1} - T_{s1}]$, where $T_{ei}$ is the duration of stimulus encoding for Task $i$ ($i = 1, 2$), $T_{si}$ is the duration of response selection for Task $i$ ($i = 1, 2$), $T_{m1}$ is the duration of movement production for Task 1 (i.e., the Task 1 MRP), and SOA is the shortest stimulus-onset asynchrony in the experiment.

change much, if at all, from Session 2 to Session 3. Thus, according to the multiple structural-bottleneck hypothesis (De Jong, 1993), either one or another of two circumstances should have prevailed: (a) During both sessions, the duration of response selection for the easy Task 2 should have been less than the Task 1 MRP, and S-R compatibility effects should have been underadditive with SOA effects throughout practice; or (b) during both sessions, the duration of response selection for the easy Task 2 should have been greater than the Task 1 MRP, and S-R compatibility effects should have been additive with SOA effects throughout practice. Yet neither of these alternative circumstances occurred in actual practice. Although the mean RT for the easy Task 2 stayed virtually constant across sessions, the effects of SOA and S-R compatibility on mean Task 2 RTs were additive during Session 2 but became underadditive during Session 3 (see Figures 7 and 8). Given this evolving pattern of null and underadditive interactions, De Jong's multiple structural-bottleneck hypothesis should be rejected. Instead, our results again suggest the involvement of flexible task-scheduling strategies based on adaptive executive control.

Nevertheless, an elaborated version of the multiple structural-bottleneck hypothesis still might have some merit. Suppose that there are not only structural response-selection and modality-specific movement-production bottlenecks but also "higher-order control processes" (De Jong, 1995) through which response selection for Task 2 can be started at various alternative moments after the completion of response selection for Task 1. Furthermore, suppose that early in practice, these control processes delay the start of Task 2 response selection until well after Task 1 movement production has begun, whereas later in practice, Task 2 response selection is started immediately after Task 1 response selection has ended. Then these joint assumptions conceivably could account for the evolving pattern of additive and underadditive effects caused by SOA and response-selection difficulty during Sessions 2 and 3 of Experiment 3. Perhaps the additive effects found during Session 2 stemmed from the start of Task 2 response selection being delayed long enough that a manual movement-production bottleneck did not contribute any temporal slack to performance of the easy Task 2 at short SOAs. In contrast, perhaps the underadditive effects found during Session 3 stemmed from Task 2 response selection being started soon enough that this bottleneck did contribute such slack.

Yet even this elaboration of the multiple structural-bottleneck hypothesis—with two types of immutable structural bottlenecks complemented by higher-order control processes—is inadequate. It may be rejected on the basis of our fourth experiment. Here we showed that when contributions from modality-specific bottlenecks in movement production are precluded, underadditive effects of SOA and response-selection difficulty still occur, as expected from our AEC models but not from extant bottleneck hypotheses. Experiment 4 also let us test further predictions about practice effects on dual-task performance.

## Experiment 4

Experiment 4 was identical to Experiment 3 except that it combined an auditory-vocal (instead of auditory-manual) Task 1 with a visual-manual Task 2. This new task combination provides a more powerful test of our AEC models versus the multiple structural-bottleneck hypothesis. According to both the original and elaborated versions of the latter hypothesis, movement-production bottlenecks are modality specific (De Jong, 1993, 1995). Consequently, if two tasks require responses in different (e.g., vocal and manual) modalities, then no movement-production bottleneck could delay them. For example, De Jong (1993) had participants perform primary and secondary tasks with manual and pedal responses, respectively, to eliminate dual-task interference caused by putative movement-production bottlenecks, and the obtained results suggested that he succeeded. Similarly, McLeod (1977; McLeod & Posner, 1984) showed that auditory-vocal and visual-manual tasks could be performed with little dual-task interference, which suggests that a structural bottleneck does not preclude concurrent movement production for them. On the basis of the multiple structural-bottleneck hypothesis, we therefore expected that during Experiment 4, only a structural cognitive RSB would interfere with performance of Task 2 under the PRP procedure. Thus, because of reasons outlined already (e.g., see Figure 1), these hypotheses predict that Experiment 4 should yield additive effects of SOA and response-selection difficulty (i.e., S-R compatibility) on mean RTs for Task 2; under present conditions, a Task 1 MRP (i.e., motor refractory period) cannot obscure this additivity. As before, however, if our AEC models are correct, then SOA and response-selection difficulty may affect mean Task 2 RTs underadditively again, because the response-selection processes for Tasks 1 and 2 still may be concurrent. Thus, observing such underadditivity here would strongly support our AEC models over the extant bottleneck hypotheses.

Our AEC models and heuristic principles of task scheduling (Meyer & Kieras, 1997a, 1997b, 1999; Meyer et al., 1995) also let us make some additional tentative predictions regarding practice effects on dual-task performance in Experiment 4. Unlike in Experiment 3, the two tasks of Experiment 4 involved different (vocal and manual) response modalities. Although Experiment 4 again included a highly incompatible S-R mapping for the difficult Task 2, this new response-modality difference between Tasks 1 and 2 might have compensatory benefits. Specifically, it could help encourage participants to adopt daring task scheduling (unpostponed Task 2 response selection) early in practice, because they would not have to be concerned about concurrently selected Task 2 responses perhaps "jamming" a modality-specific single-channel (bottleneck) mechanism that produces Task 1 response movements. Thus, we expected that Experiment 4 might yield underadditive effects of SOA and S-R compatibility on mean Task 2 RTs throughout both Sessions 2 and 3 (cf. Table 2, middle column). As we show later, this expectation was confirmed, further illustrating the flexibility of dual-task performance in

ways contrary to hypotheses that incorporate immutable structural RSBs.

## Method

### Participants

Six right-handed undergraduate students (3 men and 3 women) at the University of Michigan participated as paid volunteers. They came from the same population as those in the previous experiments but had not been tested previously. Participants had normal or corrected-to-normal vision and were paid $5.00 per hour plus a bonus based on the quality of their performance.

### Apparatus

The apparatus was the same as that used in previous experiments. Participants' vocal responses triggered an MED Associates voice-activated switch (ANL-923), which signaled the computer that a vocal response had occurred.

### Design and Procedure

All aspects of the design and procedure were identical to those in Experiment 3 except that Task 1 required vocal rather than manual responses. Participants responded by saying "low" for the 1120-Hz tone and "high" for the 1450-Hz tone.[12]

## Results

Outliers were trimmed and remaining data were analyzed in the same way as those from Experiments 2 and 3. The outlier-trimming procedure removed 5.7% of the original data set, leaving 6,882 response pairs. Figure 10 shows the mean RTs of correct responses for each task in Sessions 2 and 3 as a function of Task 2 response-selection difficulty and SOA. Table 5 shows the corresponding error rates. In what follows, we discuss the results more fully for each task.

### Task 1

*Reaction times.* The main effect of Session on Task 1 RTs was reliable, $F(1, 5) = 10.59, p < .05$. Across sessions, mean Task 1 RTs decreased from 510 to 493 ms. There was also a small but reliable main effect of SOA, $F(4, 20) = 3.51, p < .05$. Across sessions and Task 2 difficulty, Task 1 RTs increased as SOA decreased. Neither the main effect of Task 2 response-selection difficulty, $F(1, 5) = 1.85, p > .20$, nor any interaction was reliable.

*Error rates.* The overall Task 1 error rate was 2.92%. The main effect of SOA was reliable, $F(4, 20) = 10.42, p < .0005$. Participants made more errors on Task 1 when the SOA was short (4.69%) than when the SOA was long (2.34%). The interaction between the effects of session and SOA was also reliable, $F(4, 20) = 3.86, p < .05$. SOA affected Task 1 error rates more on Session 2 than on Session 3. No other main effects or interactions were reliable.

### Task 2

*Reaction times.* The main effects of Task 2 response-selection difficulty, $F(1, 5) = 52.63, p < .001$, and SOA,

$F(1, 5) = 15.79, p < .0005$, and the interaction between these effects, $F(4, 20) = 12.26, p < .0005$, on Task 2 RTs were reliable. Across sessions, mean Task 2 RTs were faster when Task 2 was easy (458 ms) than when Task 2 was hard (540 ms), and there was a moderate SOA effect. Session did not affect Task 2 RTs reliably, $F(1, 5) = 0.42$, nor did it interact with the effects of Task 2 difficulty and SOA, $F(4, 20) = 1.21, p > .30$. There were no other reliable interactions.

Figure 11 shows the response-selection difficulty effects on mean Task 2 RTs at the shortest and longest SOAs for Sessions 2 and 3. The differences between the difficulty effects at these SOAs (81 ms in Session 2 and 48 ms in Session 3) were reliable during each session: $t(5) = 4.99$, $p < .005$ for Session 2, and $t(5) = 4.77, p < .01$ for Session 3.

*Error rates.* The overall Task 2 error rate was 2.07%. No main effects or interactions on errors for Task 2 were reliable.

*Comparisons of Task 2 RTs for Experiments 3 and 4.* The mean Task 2 RTs for Experiments 3 and 4 were combined and analyzed through an ANOVA that included experiment, session, Task 2 response-selection difficulty, and SOA as factors. Here several effects were reliable. For example, there was a reliable four-way interaction among the effects of experiment, session, Task 2 response-selection difficulty, and SOA, $F(4, 48) = 3.44, p < .01$. In Experiment 4, the two-way interaction between the effects of Task 2 difficulty and SOA on mean Task 2 RTs occurred during both sessions, whereas in Experiment 3, it occurred only during Session 3. In addition, there was a reliable interaction between the effects of experiment and SOA, $F(4, 48) = 4.42, p < .005$. The overall SOA effect on mean Task 2 RTs was substantially smaller in Experiment 4 than in Experiment 3.

## Discussion

Experiment 4 yielded more evidence that dual-task performance is mediated by adaptive executive control with flexible task scheduling. We found no immutable structural RSB during this experiment. In both Sessions 2 and 3, SOA and S-R compatibility (a prototypical manipulation of response-selection difficulty) had markedly underadditive effects on mean Task 2 RTs. Contrary to multiple structural-bottleneck hypotheses (De Jong, 1993, 1995), this underadditivity occurred even though the present primary and secondary tasks involved different (i.e., vocal and manual) response modalities, which eliminated potential confounding contributions from putative Task 1 MRPs (cf. *Discussion* section of Experiment 3). Also, PRP effects (i.e., differences between mean Task 2 RTs at short and long SOAs) were much less here than in Experiment 3, even though the S-R mappings for Task 2 were the same as before and even though the mean Task 1 RTs were somewhat longer than before. The traditional RSB hypothesis cannot account for

---

[12] Out-of-order errors occurred on fewer than 1.00% of trials overall.
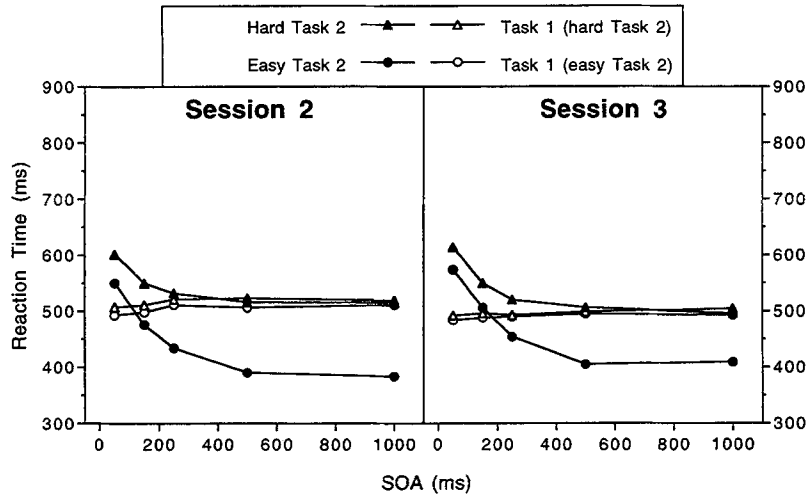
*Figure 10.* Mean reaction times for Tasks 1 and 2 as a function of Task 2 response-selection difficulty and stimulus onset asynchrony (SOA) in Sessions 2 and 3 of Experiment 4.

this latter pattern either.[13] Instead, these results accord much better with the aforementioned predictions of an AEC model whose executive processes consistently enable concurrent response-selection processes through daring task scheduling that uses a late (post-response-selection) Task 2 lockout point for obeying instructions of the PRP procedure.

Another important aspect of the results from Experiment 4 concerns changes of performance across sessions, which contrast with those from Experiment 3. In Experiment 3, SOA and response-selection difficulty affected mean Task 2 RTs additively during Session 2 but underadditively during Session 3. There it appeared that participants initially adopted a cautious scheduling strategy (postponed response selection for Task 2) and subsequently shifted to a daring scheduling strategy (unpostponed Task 2 response selection)

Table 5

*Mean Error Rates (%) in Sessions 2 and 3 of Experiment 4 for Each Task as a Function of Task 2 Response-Selection Difficulty and Stimulus Onset Asynchrony (SOA)*

| Session and Task 2 type | SOA (ms) | | | | |
|---|---|---|---|---|---|
| | 50 | 150 | 250 | 500 | 1,000 |
| | Task 1 | | | | |
| Session 2 | | | | | |
| Hard | 2.86 | 1.30 | 2.86 | 2.35 | 1.30 |
| Easy | 7.82 | 2.87 | 1.57 | 2.08 | 1.56 |
| Session 3 | | | | | |
| Hard | 3.13 | 1.30 | 3.39 | 1.82 | 3.65 |
| Easy | 4.95 | 4.95 | 2.87 | 2.87 | 2.87 |
| | Task 2 | | | | |
| Session 2 | | | | | |
| Hard | 1.30 | 0.78 | 2.08 | 1.56 | 2.60 |
| Easy | 0.52 | 1.56 | 1.82 | 1.83 | 3.39 |
| Session 3 | | | | | |
| Hard | 0.52 | 1.82 | 1.82 | 4.17 | 3.12 |
| Easy | 1.04 | 1.56 | 3.65 | 2.87 | 3.39 |

as practice progressed (cf. Table 2, middle column). However, in Experiment 4, the effects of SOA and response-selection difficulty were underadditive during both sessions, and the absolute magnitude of their interaction decreased from Session 2 to Session 3. The latter decrease occurred simply because over the course of practice, the magnitude of the difficulty effect on mean Task 2 RTs decreased more at the longest SOA than at the shortest SOA. Such a pattern suggests that Experiment 4's participants, unlike those of Experiment 3, initially and consistently adopted a daring scheduling strategy regardless of the evolution in Task 2 response selection (see Table 2, lower right cell). These differences occurred even though both experiments involved the same compatible and incompatible S-R mappings for Task 2.

As anticipated already, this is what we expected might happen on the basis of the AEC models and the heuristic principles of task scheduling. Given our expectations, conducting Experiment 4 with primary and secondary tasks that involved distinct (i.e., vocal and manual) response modalities presumably helped encourage rapid adoption and consistent use of daring task scheduling, because the response movements required for each of these tasks were programmed by distinct modality-specific movement-production mechanisms. That participants' strategy of task scheduling was indeed daring throughout Experiment 4, whereas it changed from cautious to daring in Experiment 3 when both tasks involved the same response modality, therefore con-

---

[13] For example, because Task 1 of Experiment 4 again involved two alternative S-R pairs, and because the stimuli (auditory tones) of Task 1 were the same as in Experiment 3, the RSB hypothesis implies that the size of the PRP effect should have been the same as (not substantially less than) before. The substantial decrease of the Task 2 PRP effect across experiments also suggests that an immutable structural movement-production bottleneck did not interfere with dual-task performance during Experiment 4.
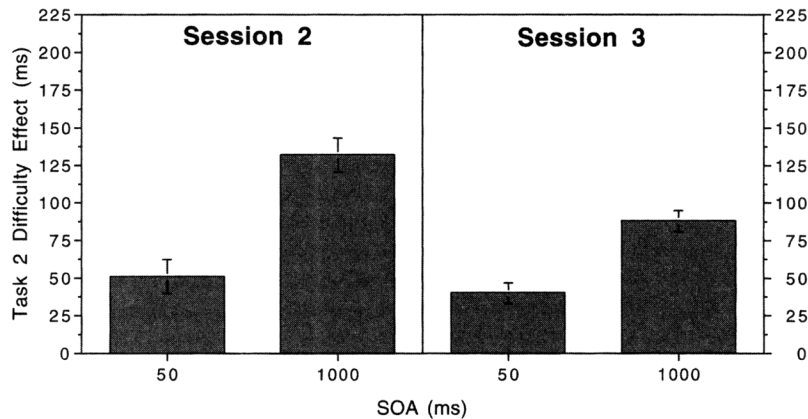
Figure 11. Response-selection difficulty effects and standard error bars for mean Task 2 reaction times at the shortest and longest stimulus onset asynchronies (SOAs) in Sessions 2 and 3 of Experiment 4. Standard error bars were derived from the mean squared error of the F test for determining the reliability of the difference between the heights of the bars.

firms more fully our AEC models and their underlying functional rationale.

## General Discussion

The present four experiments show that RSBs are neither structural nor immutable but rather that dual-task performance is mediated by potentially concurrent response-selection processes, adaptive executive control, and flexible strategies of task scheduling. In Experiments 1 and 2 we replicated and extended studies by some previous investigators with the PRP procedure in which the secondary tasks involved a manipulation of S-R numerosity (cf. Hawkins et al., 1979; Karlin & Kestenbaum, 1968). Like them, we found consistent underadditive effects of SOA and S-R numerosity on mean Task 2 RTs. Contrary to arguments by Pashler (1994a), such underadditivity occurred even when Task 2 was always a choice-reaction task in which the stimuli were familiar and highly discriminable printed symbols. Our results cannot be attributed merely to idiosyncratic characteristics of simple-reaction secondary tasks or to confounded effects of S-R numerosity on stimulus encoding. Instead, it appears that our results stemmed from participants' use of a daring task-scheduling strategy with concurrent response-selection processes for Tasks 1 and 2 at short SOAs. In Experiment 3, this finding was generalized. After participants had some practice at dual-task performance there, mean Task 2 RTs were affected underadditively by SOA and S-R compatibility, a prototypical factor for influencing secondary-task response selection and manifesting its temporal overlap with primary-task response selection. Experiment 4 further revealed that such underadditivity occurred not only when the primary and secondary tasks both involved manual responses but also when they involved different (viz., vocal and manual) response modalities. With the latter result, the traditional RSB hypothesis and various versions of a multiple structural-bottleneck hypothesis (De Jong, 1993, 1995) may be rejected in favor of AEC models

(Meyer & Kieras, 1997a, 1997b, 1999; Meyer et al., 1995). Moreover, unlike these former deficient hypotheses, the AEC models provide a principled account—based on heuristic principles of task scheduling—for different patterns of practice effects observed across our experiments as a result of systematic changes in the prevailing S-R mappings and response-modality combinations.

### Explanation of Positive Response-Selection Difficulty Effects at Short SOAs

Given that there were concurrent response-selection processes for Tasks 1 and 2 in our experiments, one might wonder why response-selection difficulty (i.e., S-R numerosity and S-R compatibility) had any positive effects on mean Task 2 RTs at the shortest SOAs. Should not temporal slack after Task 2 response selection but before Task 2 movement production have made these effects strictly null instead? Again, an answer is provided by the present theoretical framework (see Figure 3).

Because of between-trial variability in the processing-stage durations for Tasks 1 and 2 of the PRP procedure, and because of technical requirements associated with unlocking Task 2 after its movement-production stage has been postponed, the executive processes of some AEC models enable concurrent response-selection processes at the start of each trial but later suspend Task 2 response selection briefly on a subset of trials while Task 2 is being unlocked. In essence, this brief intermediate suspension of Task 2 response selection constitutes a temporary strategic "bottleneck" that can yield a residual positive effect of response-selection difficulty on mean Task 2 RTs at short SOAs (see Meyer & Kieras, 1997a, for a more detailed explanation). However, the source of this effect is neither structural nor immutable, unlike what the RSB and multiple structural-bottleneck hypotheses assume. Thus, the modest positive difficulty effects on mean Task 2 RTs that we observed at short SOAs provide no consolation for these hypotheses.

## Explanation of Previous Additive SOA and Response-Selection Difficulty Effects

Another remaining question concerns why additive rather than underadditive effects of SOA and response-selection difficulty on mean Task 2 RTs have occurred in some previous studies with the PRP procedure (Carrier & Pashler, 1995; Fagot & Pashler, 1992; McCann & Johnston, 1992; Pashler, 1984; Pashler & Johnston, 1989; Ruthruff, Miller, & Lachmann, 1995; Van Selst & Jolicoeur, 1997). Although the present experiments offer no definitive answers, our results and AEC models suggest some interesting possibilities. For example, as mentioned earlier and shown in Experiment 3, participants may need multiple sessions of practice to adopt relatively efficient (daring) strategies of task scheduling that entail concurrent response-selection processes. Unfortunately, many previous PRP studies included only one session (e.g., Carrier & Pashler, 1995; Fagot & Pashler, 1992; Pashler & Johnston, 1989; Ruthruff, Miller, & Lachmann, 1995; Van Selst & Jolicoeur, 1997, Exps. 1, 3, & 4). Perhaps they would have yielded more underadditive effects if their participants received additional practice.

In previous studies with multiple practice sessions, the experimental designs had other features from which additive effects of SOA and response-selection difficulty might stem. For example, some of them included participants who were highly familiar with the rationale and favorite hypotheses being tested through the PRP procedure (Van Selst & Jolicoeur, 1997, Experiment 2), which could have biased the obtained results toward additivity. Also, for some of them, the longest SOA was less than the mean Task 1 RT (Fagot & Pashler, 1992; Pashler & Johnston, 1989; Ruthruff, Miller, & Lachmann, 1995), which could have reduced the statistical power to detect an underadditive interaction between SOA and response-selection difficulty effects on mean Task 2 RTs. Specifically, if Task 1 processing stages are still underway at the longest SOA, then temporal slack before Task 2 movement production can absorb at least some of the difficulty effect throughout the entire prevailing SOA range, thereby attenuating the observed SOA-by-difficulty interaction considerably. The latter concern is especially problematic because the effects of response-selection difficulty on mean Task 2 RTs at the longest SOAs have been small (viz., 50 ms or less) in some previous PRP studies (e.g., Pashler, 1984; Pashler & Johnston, 1989). With such small effects, the power to discover interactions involving them is even more drastically limited.

Yet these various critical considerations do not entirely disarm one previous PRP study that yielded putative support for the traditional RSB hypothesis. This study was conducted by McCann and Johnston (1992). During two experiments that each involved two sessions of practice, they manipulated response-selection difficulty (viz., S-R compatibility) for Task 2, obtaining moderate (55 to 72 ms), approximately additive, difficulty effects on mean Task 2 RTs across a wide range of SOAs (from 50 to 800 ms). Nevertheless, consistent with our AEC models, there are some further plausible explanations for why McCann and Johnston failed to find reliable underadditive effects of SOA and response-selection difficulty.

On the second (and last) session of McCann and Johnston's (1992) first experiment, the mean RT for even the easy version of Task 2 exceeded 600 ms at the longest SOA. The secondary task required participants to base their responses on both the shape and size of a geometric visual stimulus. As we found from our own Experiment 3 (cf. Figure 8), perhaps these requirements induced participants initially to adopt a cautious task-scheduling strategy, and perhaps beyond their second session of practice, they would subsequently have shifted to a more daring scheduling strategy with concurrent response-selection processes. Indeed, that this could be so is hinted at by details of the available RT data from Session 2 of McCann and Johnston's first experiment. There the response-selection difficulty effect on mean Task 2 RTs was 17 ms less at the shortest SOA than at the longest SOA, and this difference was nearly "significant" by conventional standards, $F(3, 69) = 1.94$, $p < .075$ (one tailed). Such marginal underadditivity is what one might expect if a shift from cautious to daring task scheduling began during the latter part of Session 2 (cf. Table 2).

Unfortunately, McCann and Johnston's (1992) second experiment provides no additional relevant tests of these possibilities. It involved less complex S-R mappings for Task 2 than did their first experiment, but on each trial after the Task 2 stimulus appeared, participants had to make eye movements before they could identify it. This requirement, which stemmed from the nature of the visual displays used there, added at least 150 ms of "dead time" on average to the shortest SOA (cf. Abrams & Jonides, 1988). As a result, the power of McCann and Johnston's second experiment to reveal concurrent response-selection processes at the shortest SOA was greatly diminished, and the data from it are entirely consistent with our AEC models (Meyer & Kieras, 1997b, pp. 762–765).

## Explanation of Other PRP Phenomena

Of course, there are many other phenomena from previous PRP studies that merit explanation as well. Although the effects of SOA and response-selection difficulty on mean Task 2 RTs may become underadditive as dual-task practice progresses, some investigators have found that PRP effects (elevation of mean Task 2 RTs at short SOAs) persist over many practice sessions (Gottsdanker & Stelmach, 1971; cf. Van Selst et al., 1997). Performing primary tasks that do not require overt responses still induces PRP effects in secondary tasks (Davis, 1959; Fraisse, 1957; Nickerson, 1965; Van Selst & Johnston, 1997). PRP effects are observed not only in neurologically intact participants but also in split-brain patients, who supposedly have separated cerebral hemispheres for performing primary and secondary tasks independently (Ivry et al., 1998; Pashler et al., 1994; cf. Pashler & O'Brien, 1993). Even when participants are encouraged to perform two tasks with equal priority, substantial decrements in dual-task performance sometimes occur (e.g., Pashler, 1994b; Ruthruff, Pashler, & Klaasen, 1995). However, emphatic equal-priority instructions, together with

other experimental-design features that eliminate artifactual dependencies between tasks, can dramatically reduce or eliminate these decrements, yielding virtually perfect time-sharing (Allport, Antonis, & Reynolds, 1972; Greenwald & Shulman, 1973; Koch, 1993, 1994; Schumacher et al., 1997). Also, dual-task performance decrements may be attenuated substantially under the PRP procedure despite its unequal-priority instructions, if the primary task does not require overt responses (Davis, 1959; Fraisse, 1957; Kay & Weiss, 1961; Nickerson, 1965; Van Selst & Johnston, 1997) or if the same SOA is used throughout an entire trial block (Borger, 1963). This complex pattern of phenomena therefore poses a strong challenge to any prospective theory of dual-task performance.

We believe that our AEC models and their underlying theoretical framework, the executive-process interactive control (EPIC) information-processing architecture (Meyer & Kieras, 1997a, 1997b, 1999; Meyer et al., 1995), may meet this challenge successfully. As Meyer and Kieras (1997b, pp. 772–775) discussed, AEC models and EPIC provide natural parsimonious accounts for all of the afore-mentioned PRP phenomena. Essentially, these phenomena as well as those that occur in related dual-task paradigms (e.g., Ballas et al., 1992) can be explained—and sometimes predicted on a priori grounds—in terms of joint relationships among prevailing instructions about relative task priorities, AEC processes, flexible task-scheduling strategies, and constraints on peripheral perceptual-motor mechanisms. For any particular phenomenon to be understood properly, a theorist must encompass all of these facets of dual-task performance.

### Directions for Future Research
### on Dual-Task Performance

From the perspective of our AEC models and EPIC architecture, there are many promising directions for future research on dual-task performance. We have found that dual-task performance is mediated by executive processes that use flexible strategies of task scheduling in order to satisfy prevailing instructions about relative task priorities. The present four experiments offer some information about what factors contribute to the choice of a particular strategy. For example, difficult secondary tasks whose S-R mappings are incompatible tend to encourage initial adoption of cautious task scheduling with temporary postponement of secondary-task response selection (Experiment 3). Such caution and response-selection postponement sometimes happen although, in principle, concurrent response-selection processes for the primary and secondary tasks would be possible (cf. Experiments 1, 2, and 4). Nevertheless, even when task scheduling at first is cautious, moderate amounts of dual-task practice may lead to the adoption of more daring scheduling strategies (Experiment 3). Daring task scheduling also appears to be encouraged by combinations of primary and secondary tasks that involve different (e.g., vocal and manual) response modalities (Experiment 4). Nevertheless, much remains to be discovered about what factors govern the choice of particular task scheduling

strategies and how executive processes implement them adaptively.

The need for such discoveries raises fundamental questions. Do small nuances of instructions about relative task priorities influence which type of scheduling strategy is chosen? Can special regimens of dual-task training facilitate the adoption of efficient (e.g., daring) task scheduling (cf. Gopher, 1993; Meyer et al., 1995)? Is the choice of a particular scheduling strategy determined in significant part by people's individual differences and personality characteristics (cf. Dickman & Meyer, 1988; Meyer et al., 1995)? Some additional research with the PRP procedure and other related dual-task paradigms has begun to provide further tentative positive answers to these and other related questions (Meyer & Kieras, 1997b, 1999; Meyer et al., 1995; Schumacher et al., 1997). Yet more work will be essential for obtaining definitive answers and for understanding exactly what kinds of adaptive executive control underlie them.

## References

Abrams, R. A., & Jonides, J. (1988). Programming saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 428–443.

Allport, D. A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single-channel hypothesis. *Quarterly Journal of Experimental Psychology, 24,* 225–235.

Ballas, J. A., Heitmeyer, C. L., & Perez, M. A. (1992). *Direct manipulation and intermittent automation in advanced cockpits* (Tech. Rep. NRL/FR/5534-92-9375). Washington, DC: Naval Research Laboratory.

Borger, R. (1963). The refractory period and serial choice reactions. *Quarterly Journal of Experimental Psychology, 15,* 1–12.

Brainard, R. W., Irby, T. S., Fitts, P. M., & Alluisi, E. A. (1962). Some variables influencing the rate of gain of information. *Journal of Experimental Psychology, 63,* 105–110.

Broadbent, D. E., & Gregory, M. (1965). On the interaction of S-R compatibility with other variables affecting reaction time. *British Journal of Psychology, 56,* 61–67.

Brown, T. L., & Carr, T. H. (1989). Automaticity in skill acquisition: Mechanisms for reading interference in concurrent performance. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 686–700.

Carrier, M. L., & Pashler, H. (1995). Attentional limits in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1339–1348.

Davis, R. (1959). The role of "attention" in the psychological refractory period. *Quarterly Journal of Experimental Psychology, 11,* 211–220.

De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 965–980.

De Jong, R. (1995). The role of preparation in overlapping-task performance. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 48A,* 22–25.

Dickman, S. J., & Meyer, D. E. (1988). Impulsivity and speed–accuracy tradeoffs in information processing. *Journal of Personality and Social Psychology, 54,* 274–290.

Duncan, J. (1977). Response selection errors in spatial choice reaction tasks. *Quarterly Journal of Experimental Psychology, 29,* 415–423.

Fagot, C., & Pashler, H. (1992). Making two responses to a single object: Implications for the central attentional bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1058–1079.

Fitts, P. M., & Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology, 46,* 199–210.

Fraisse, P. (1957). La periode refractaire psychologique. *Annee Psychologique, 57,* 315–328.

Gopher, D. (1993). Attentional control: Acquisition and execution of attentional strategies. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV* (pp. 299–322). Cambridge, MA: MIT Press.

Gottsdanker, R., & Stelmach, G. E. (1971). The persistence of psychological refractoriness. *Journal of Motor Behavior, 3,* 301–312.

Greenwald, A. G., & Shulman, H. (1973). On doing two things at once II: Elimination of the psychological refractory period. *Journal of Experimental Psychology, 101,* 70–76.

Hawkins, H. L., Rodriquez, E., & Reicher, G. M. (1979). *Is time-sharing a general ability?* (Tech. Rep. 3). Eugene: University of Oregon.

Ivry, R. B., Franz, E. A., Kingstone, A., & Johnston, J. C. (1998). The PRP effect following callosotomy: Uncoupling of lateralized response codes. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 463–480.

Karlin, L., & Kestenbaum, R. (1968). Effects of number of alternatives on the psychological refractory period. *Quarterly Journal of Experimental Psychology, 20,* 167–178.

Kay, H., & Weiss, A. D. (1961). Relationship between simple and serial reaction time. *Nature, 191,* 790–791.

Keele, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin, 70,* 387–403.

Keele, S. W. (1973). *Attention and human performance.* Pacific Palisades, CA: Goodyear.

Koch, R. (1993). *Die psychologische Refraktärperiode.* Unpublished doctoral dissertation, University of Bielefeld, Bielefeld, Germany.

Koch, R. (1994, December). *Hick's law and the psychological refractory period.* Paper presented at the KNAW Symposium on Discrete Versus Continuous Information Processing, Amsterdam, The Netherlands.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus–response compatibility—A model and taxonomy. *Psychological Review, 97,* 253–270.

McCann, R. S., & Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 471–484.

McLeod, P. (1977). A dual task response modality effect: Support for multiprocessor models of attention. *Quarterly Journal of Experimental Psychology, 29,* 651–667.

McLeod, P., & Posner, M. I. (1984). Privileged loops from percept to act. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X* (pp. 55–66). Hillsdale, NJ: Erlbaum.

Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review, 104,* 3–65.

Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory-period phenomena. *Psychological Review, 104,* 749–791.

Meyer, D. E., & Kieras, D. E. (1999). Précis to a practical unified theory of cognition and action: Some lessons from computa-

tional modeling of human multiple-task performance. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII* (pp. 15–88). Cambridge, MA: MIT Press.

Meyer, D. E., Kieras, D. E., Lauber, E., Schumacher, E. H., Glass, J., Zurbriggen, E., Gmeindl, L., & Apfelblat, D. (1995). Adaptive executive control: Flexible multiple-task performance without pervasive immutable response-selection bottlenecks. *Acta Psychologica, 90,* 163–190.

Miller, J., & Ulrich, R. (1998). Locus of the effect of the number of alternative responses: Evidence from the lateralized readiness potential. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1215–1231.

Morin, R. E., & Forrin, B. (1965). Information-processing: Choice reaction times of first- and third-grade students for two types of associations. *Child Development, 36,* 713–720.

Nickerson, R. S. (1965). Response time to the second of two successive signals as a function of absolute and relative duration of intersignal interval. *Perceptual and Motor Skills, 21,* 3–10.

Pashler, H. (1984). Processing stages in overlapping tasks. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 358–377.

Pashler, H. (1994a). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116,* 220–244.

Pashler, H. (1994b). Graded capacity-sharing in dual-task interference? *Journal of Experimental Psychology: Human Perception and Performance, 20,* 330–342.

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

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

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

Pashler, H., Luck, S. J., Hillyard, S. A., Mangun, G. R., O'Brien, S., & Gazzaniga, M. S. (1994). Sequential operation of the disconnected cerebral hemispheres in "split-brain" patients. *Neuroreport, 5,* 2381–2384.

Pashler, H., & O'Brien, S. (1993). Dual-task interference and the cerebral hemispheres. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 315–330.

Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General, 109,* 444–474.

Ruthruff, E., Miller, J., & Lachmann, T. (1995). Does mental rotation require central mechanisms? *Journal of Experimental Psychology: Human Perception and Performance, 21,* 552–570.

Ruthruff, E., Pashler, H., & Klaasen, A. (1995, November). *Preparation and strategy in dual-task interference.* Poster presented at the meeting of the Psychonomic Society, Los Angeles, CA.

Sanders, A. F. (1980). Stage analysis of reaction processes. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 331–354). New York: North-Holland.

Schumacher, E. H., Glass, J., Lauber, E. J., Gmeindl, L., Woodside, B. J., Kieras, D. E., & Meyer, D. E. (1996, November). *Concurrent response-selection processes in multiple-task performance.* Poster presented at the meeting of the Psychonomic Society, Chicago, IL.

Schumacher, E. H., Seymour, T., Glass, J. M., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (1997, November). *Virtually perfect time sharing in dual-task performance.* Paper presented at the meeting of the Psychonomic Society, Philadelphia, PA.

Schweickert, R., & Boggs, G. J. (1984). Models of central capacity and concurrency. *Journal of Mathematical Psychology, 28,* 223–281.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. In W. G. Koster (Ed.), *Attention and performance II* (pp. 276–315). Amsterdam: North-Holland.

Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science: Vol. 4. Methods, models, and conceptual issues* (pp. 703–863). Cambridge, MA: MIT Press.

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review, 61,* 401–409.

Theios, J. (1973). Reaction time measurements in the study of memory processes: Theory and data. In G. H. Bower (Ed.), *The psychology of learning and motivation VII* (pp. 43–85). New York: Academic Press.

Van Selst, M., & Johnston, J. C. (1997, November). *Response modality modulates PRP interference, even when task 1 is a no-go trial.* Poster presented at the meeting of the Psychonomic Society, Philadelphia, PA.

Van Selst, M., & Jolicoeur, P. (1997). Decision and response in dual-task interference. *Cognitive Psychology, 33,* 266–307.

Van Selst, M., Ruthruff, E., & Johnston, J. C. (1997, November). *Practice drastically reduces interference in the psychological refractory period paradigm.* Paper presented at the meeting of the Psychonomic Society, Philadelphia, PA.

Welford, A. T. (1980). The single-channel hypothesis. In A. T. Welford (Ed.), *Reaction times* (pp. 215–252). New York: Academic Press.