Journal of Experimental Psychology:
Human Perception and Performance
2001, Vol. 27, No. 1, 3–21

DOI: 10.1037//0096-1523.27.1.3

Why Practice Reduces Dual-Task Interference

Eric Ruthruff and James C. Johnston
National Aeronautics and Space Administration
Ames Research Center

Mark Van Selst
San Jose State University

M. A. Van Selst, E. Ruthruff, and J. C. Johnston (1999) found that practice dramatically reduced dual-task interference in a Psychological Refractory Period (PRP) paradigm with 1 vocal response and 1 manual response. Results from 3 further experiments using the highly trained participants of M. A. Van Selst et al. (1999) support 4 main conclusions: (a) A processing bottleneck exists even after extensive practice; (b) the principal cause of the reduction in PRP interference with practice is shortening of Task 1 bottleneck stages; (c) a secondary cause is that 1 or more, but not all, of the Task 2 substages that are postponed before practice are not postponed after practice (i.e., become automatized); and (d) the extent of PRP reduction with practice depends on the modalities of the 2 responses. A control experiment with 2 manual response tasks showed less PRP reduction with practice than that found by Van Selst et al.

Humans often have great difficulty performing two tasks at once. This difficulty has been extensively studied using the psychological refractory period (PRP) paradigm, where two stimuli—each requiring a separate speeded response—are presented in rapid succession. Typically, responses to the first stimulus are unimpaired, but responses to the second stimulus are slowed by 300 ms or more. According to the central bottleneck model, this second-task slowing reflects an inability to perform central mental operations (e.g., those involving decision-making or memory retrieval) on Task 2 while central operations on Task 1 are still underway (cf. Pashler, 1984; Pashler & Johnston, 1989; Welford, 1952, 1980). This model is well supported by a wide variety of PRP experiments (see Pashler & Johnston, 1998, for a review). One limitation of these previous experiments, however, is that they have generally used relatively unpracticed tasks. Consequently, it is unclear whether the central bottleneck model also applies to highly practiced tasks, which are commonplace in many real-world domains (e.g., aviation, manufacturing, playing a musical instrument).

To address this important issue, Van Selst, Ruthruff, and Johnston (1999) recently studied the effects of 36 practice sessions on dual-task interference. Using a PRP design with a Task 1 vocal response and a Task 2 manual response, they found that practice drastically reduced dual-task interference. On the other hand, a small residual PRP effect did remain. The pattern of factor interactions (discussed later) indicated that this residual PRP effect was due to a processing bottleneck. Van Selst et al. proposed that practice does not eliminate the bottleneck, but it does shorten the Task 1 stages that cause the bottleneck. The present article reports three new experiments that further explore the dramatic effects of practice on dual-task interference.

Background

A variety of experimental paradigms have been used to study dual-task performance. One common approach has been to measure accuracy on two continuous tasks (e.g., reading and shadowing), performed either together or alone. Several studies using this approach have found that participants can perform two tasks together nearly as accurately as they can perform the tasks alone (e.g., Allport, Antonis, & Reynolds, 1972; Hirst, Spelke, Reaves, Caharack, & Neisser, 1980; Shaffer, 1975). It is tempting to conclude, therefore, that the mental operations required by the two tasks can function simultaneously with no interference. This conclusion is unwarranted, however. It is possible that certain critical mental operations on each task do indeed interfere with one another but without causing a drop in accuracy (see Broadbent, 1982; McCann & Johnston, 1992; Pashler, 1998; Pashler & Johnston, 1998; Shaffer, 1975). For example, participants might be able to buffer the relevant perceptual or response codes for one task while temporarily working on the other. Once they complete the critical operations on one task, they might then retrieve the buffered information from the other task and simply pick up where they left off. Provided that participants can alternate back and forth between the critical operations of the two tasks before the buffered information is lost, it might be possible to perform both tasks without error. Thus, paradigms that rely on accuracy data might conceal the fact that certain critical operations of the two tasks interfere with one another.

To reliably detect processing delays caused by dual-task interference, it is necessary to measure response time (RT) to both tasks. RT measurement, of course, requires that both the presentation of stimuli and the execution of responses be precisely timed. One design that meets this requirement is the PRP design. In fact, this design is so well suited to measuring processing delays that it has become the primary tool for assessing theories of dual-task performance (see Pashler & Johnston, 1998).

Eric Ruthruff and James C. Johnston, National Aeronautics and Space Administration (NASA) Ames Research Center, Moffett Field, California; Mark Van Selst, Department of Psychology, San Jose State University.

This research was supported by grants from the National Research Council and the National Aeronautics and Space Administration. We thank Tom Carr, Rich Ivry, Valerie Lewis, Rob McCann, Hal Pashler, Roger Remington, and Carlo Umilta for useful commentary.

Correspondence concerning this article should be addressed to Eric Ruthruff, NASA Ames Research Center, MS 262-4, Moffett Field, California 94035. Electronic mail may be sent to eruthruff@mail.arc.nasa.gov.

Psychological Refractory Period Design

In a PRP experiment, participants perform two separate tasks, each of which has a discrete stimulus and a discrete response. The key independent variable is the temporal separation between the onset of the two stimuli, known as the *stimulus onset asynchrony* (SOA). At long SOAs (the baseline condition) the two tasks are performed essentially one at a time, whereas at short SOAs there is a high degree of task overlap. In modern PRP experiments, the SOA is varied randomly within a block of trials (rather than between blocks of trials) to ensure that participants achieve the same preparatory state prior to each SOA condition. The main question is whether the requirement to work on both tasks at the same time will prolong RT. In other words, will participants respond more slowly at short SOAs than at long SOAs?

What the vast majority of PRP experiments have found is that Task 1 responses depend very little on the SOA, but Task 2 responses slow dramatically as the SOA becomes shorter. This Task 2 slowing is known as the *Psychological Refractory Period effect,* or *PRP effect* for short. Very large PRP effects (300+ ms) have been observed using a wide variety of judgments and a wide variety of input and output modalities. In fact, the PRP effect has thus far been reported to be small or nonexistent only in rare cases, all of which appear to involve tasks with extremely compatible stimulus-to-response (S-R) mappings (Greenwald, 1972; Greenwald & Shulman, 1973; Halliday, Kerr, & Elithorn, 1959; Johnston & Delgado, 1993; Pashler, Carrier, & Hoffman, 1993).

Central Bottleneck Model

Welford (1952), who noted that the PRP effect does not depend on any obvious input or output conflicts, proposed that the effect is due to an inability to perform central operations on more than one task at a time. This proposal has become known as the *central bottleneck model.* Figure 1 shows a generalized version of this model in which each task is decomposed into three processing stages[1]: a prebottleneck stage (A), a bottleneck stage (B), and a postbottleneck stage (C). By hypothesis, Stages A and C can proceed in parallel with any stage on another task. Stage B, however, is the bottleneck stage: While Stage 1B is being carried out, Stage 2B must wait. It is primarily this waiting time, or *bottleneck delay* (represented by the horizontal dashed line in Figure 1), that produces the PRP effect.

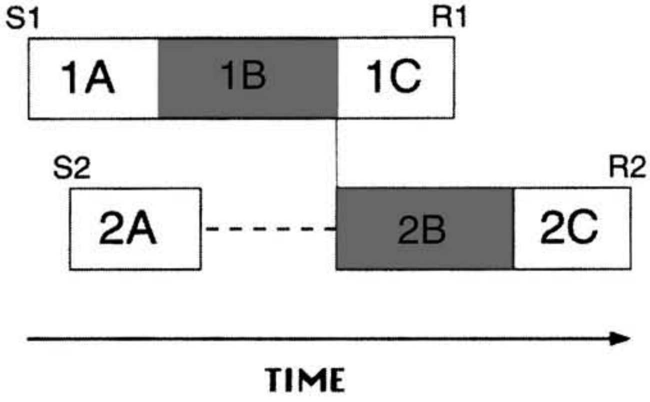

Figure 1. Generalized bottleneck model. S1 and S2 represent the stimulus onsets for Task 1 and Task 2. R1 and R2 represent the execution of the Task 1 and Task 2 responses. Processing on each task is divided into three stages, arbitrarily labeled *A, B,* and *C* so as not to presuppose exactly which mental operations are accomplished by each stage. Stage B is the bottleneck stage, meaning that while Stage 1B is underway, Stage 2B must wait.

The central bottleneck model is very simple and very general, but it makes a number of strong predictions. For example, suppose that a certain experimental manipulation increases the duration of a bottleneck or prebottleneck stage of Task 1 (i.e., Stage 1A or 1B). This manipulation should increase Task 1 response time (RT1) and, at short SOAs, this effect should carry over fully onto Task 2 response time (RT2) as well (Smith, 1969; Van Selst & Johnston, 1997). We refer to this as the *Task 1 carryover prediction.* Carryover occurs whenever there is a bottleneck delay (i.e., at short SOAs), because any delay in the completion of Stage 1B will delay the onset of Stage 2B, which will in turn delay RT2. In fact, provided there is a bottleneck delay on every trial at short SOAs, the experimental manipulation should delay RT2 by the same amount that it delays RT1. At long SOAs, meanwhile, there is no bottleneck delay and hence no carryover.

The central bottleneck model predicts a different pattern of results when an experimental manipulation increases the duration of a prebottleneck stage of Task 2 (i.e., Stage 2A). At long SOAs, RT2 is simply the sum of the times of the component stages. Thus, an increase of k ms to the duration of Stage 2A will simply add k ms to RT2. At short SOAs, however, an increase of k ms to the duration of Stage 2A will have less than a k-ms effect on RT2. The reason is that Stage 2B generally cannot begin when Stage 2A concludes but instead must wait for Stage 1B to finish (see Figure 1). Hence, a small increase in the duration of Stage 2A should not delay the onset of Stage 2B. Put another way, the bottleneck delay creates slack in the processing of Task 2, which can absorb the time added to Stage 2A (Schweickert & Boggs, 1984). In fact, if the slack time is always greater than k ms, then an increase of k ms in the duration of Stage 2A should be absorbed completely. We refer to this as the *Task 2 absorption prediction.* For a more detailed discussion of these and other predictions, see Pashler and Johnston (1998) or Schweickert and Boggs (1984).

These two predictions for the short SOA are counterintuitive in that RT2 is expected to depend strongly on the duration of prebottleneck stages of Task 1 but not on the duration of prebottleneck stages of Task 2 itself. Nevertheless, these predictions have been confirmed in a number of recent studies. The Task 1 carryover prediction was verified by McCann and Johnston (1989) and Van Selst and Johnston (1997). The Task 2 absorption prediction has been confirmed for several manipulations of early Task 2 stages, such as stimulus contrast (De Jong, 1993; Pashler, 1984; Pashler & Johnston, 1989) and categorization difficulty (Van Selst & Johnston, 1997). For a thorough review of the evidence supporting bottleneck models, see Pashler and Johnston (1998).

Locus of the Bottleneck

The abstract central bottleneck model implies that the bottleneck stages are neither the first nor the last stages in a task, but it

---

[1] Naturally, each of these three "superstages" might consist of several distinct processes, or substages.

deliberately does not specify exactly where the bottleneck occurs. However, the locus of the bottleneck can be investigated by manipulating the durations of various Task 2 stages and then determining whether the effects on RT2 are absorbed into cognitive slack (i.e., become smaller at short SOAs; see McCann & Johnston, 1992). A finding of absorption into slack supports the conclusion that the manipulated Task 2 stage occurs prior to the bottleneck. A failure to find absorption into slack supports the conclusion that the Task 2 stage occurs at or after the bottleneck.

As noted previously, Task 2 stimulus contrast manipulations and Task 2 stimulus categorization manipulations tend to be absorbed into slack (e.g., McCann & Johnston, 1989; Van Selst & Johnston, 1997), indicating that early stimulus processing, at least until the stage of character identification, occurs before the bottleneck. However, manipulations of later stages, such as decision making, response selection, mental rotation, and memory retrieval, tend to show little or no absorption (Carrier & Pashler, 1995; McCann & Johnston, 1992; Pashler, 1984; Pashler & Johnston, 1989; Ruthruff, Miller, & Lachmann, 1995; Ruthruff, Pashler, & Klaassen, in press). The most likely locus of the bottleneck, therefore, would be after stimulus identification (at least for highly familiar character stimuli) but at or before more central processing stages.

Can the Central Bottleneck Be Bypassed?

As noted earlier, very little dual-task interference occurs in certain cases where one or both tasks involve an extremely compatible, or *ideomotor compatible,* mapping of stimuli to responses (e.g., Greenwald & Shulman, 1973; Halliday et al., 1959; Johnston & Delgado, 1993; Pashler et al., 1993). For example, Greenwald and Shulman (1973) found that participants could "shadow" an auditory stimulus (i.e., repeat back what they heard) while they manually moved a switch in the direction of a visual stimulus. Presumably, ideomotor compatible tasks produce negligible PRP effects, because the stimulus codes can serve directly as response codes and thus the tasks do not require any of the central operations that constitute the bottleneck (Greenwald, 1972).

Although the bottleneck can apparently be bypassed when one or both tasks are ideomotor compatible, other attempts to bypass the bottleneck have generally failed. For example, Ruthruff, Pashler, and Klaassen (in press) strongly encouraged participants to bypass the bottleneck by (a) explicitly encouraging processing overlap, (b) placing equal emphasis on each task and asking participants to group their responses, (c) minimizing input and output conflicts, (d) training participants for three sessions, and (e) always presenting both stimuli at the same time (i.e., rather than using a variable SOA, as in the typical PRP design, the SOA was always equal to zero). Despite these efforts, large interference effects remained. In fact, the interference effects were about as large as those found using the traditional PRP paradigm. Apparently, participants cannot bypass the central bottleneck by sheer effort of will. Instead, the bottleneck (at least for relatively unpracticed tasks) appears to reflect a structural limitation inherent in the cognitive architecture (see Meyer & Kieras, 1997a, 1997b, for an opposing view).

Can extensive practice eliminate the central bottleneck? Practice often dramatically reduces RT in single-task conditions (see Logan, 1988; Pashler & Baylis, 1991), so it would not be surprising *if* practice eliminated, or at least dramatically reduced, dual-task interference as well. The central bottleneck model, in fact, clearly predicts that the PRP effect should decrease as Task 1 stage durations (1A and 1B) decrease. In addition, participants might learn to efficiently interleave the two tasks or to treat them as a single, conjoint task. Furthermore, the operations that formerly comprised the bottleneck might become automatized with practice, allowing participants to completely bypass the central bottleneck (for relevant discussion, see Bargh, 1992; Brown & Carr, 1989; Hirst et al., 1980; Logan, 1988).

It is perhaps surprising, therefore, that previous PRP studies have found relatively little effect of practice on dual-task interference (Bertelson & Tisseyre, 1969; Borger, 1963; Davis, 1956; Dutta & Walker, 1995; Halliday et al., 1959; Karlin & Kestenbaum, 1968; Van Selst & Jolicoeur, 1997). Virtually all of these studies showed a residual PRP effect of 200 ms or more after practice. It is important to note, however, that these studies all required manual responses to both tasks, which might have induced conflicts in response production. There is evidence, in fact, that manual-manual designs cause a response initiation bottleneck in addition to the central bottleneck (De Jong, 1993; see also Keele, 1973). Furthermore, the similarity of response codes in manual-manual designs might increase interference, or cross-talk, between the two response selection processes, making it difficult or impossible for them to operate concurrently. Whatever the cause of the extra interference in manual-manual designs, it seems plausible that the interference would be especially resistant to practice. Thus, although previous studies found that practice does not greatly reduce PRP interference in manual-manual designs, they leave open the question of what effect practice would have *if* response conflicts were minimized (i.e., by using tasks with distinct response modalities).

Van Selst et al. (1999)

To determine if practice can reduce or eliminate PRP interference when response conflicts are minimized, Van Selst et al. (1999)—hereinafter referred to as *VRJ*—recently studied the effects of practice in a design with one vocal response and one manual response. Task 1 required a vocal response ("high" or "low") to a tone that was high or low in pitch, while Task 2 required a manual keypress to an alphanumeric character. Thus, the tasks used different input and output modalities, which presumably served to minimize input and output conflicts (see Shaffer, 1975). VRJ also attempted to selectively manipulate the durations of several different stages on Task 1 and Task 2. This allowed them to evaluate the Task 1 carryover and Task 2 absorption predictions after practice and thus determine if the residual PRP effects, if any, were due to a processing bottleneck.

What VRJ found was a dramatic reduction in the size of the PRP effect with practice. The PRP effect in the first session was 353 ms, which is typical of a PRP experiment with relatively unpracticed tasks; but by the 18th session, the mean PRP effect was only 40 ms. Thus, practice reduced the size of the PRP effect by nearly 90%. VRJ concluded that dramatic PRP reduction is indeed possible, provided that response conflicts have been minimized. It is important to note that the S-R mappings used by VRJ were only moderately compatible; in fact, half of the Task 2 character stimuli were mapped incompatibly onto the response keys. Thus, this finding appears to be the first reported case of small PRP effects obtained when neither task was ideomotor compatible.

Although practice greatly reduced the PRP effect, it did not eliminate the PRP effect entirely.[2] Furthermore, VRJ were able to confirm two key predictions of the central bottleneck model (Task 1 carryover and Task 2 absorption predictions) after practice. These findings support the conclusion that the residual PRP effects, like the initial PRP effects, were due to a processing bottleneck. VRJ concluded, therefore, that practice shortened the durations of stages on the two tasks but did not alter the nature of the bottleneck. They referred to this straightforward extension of the bottleneck model as the *bottleneck model with stage shortening* (BSS).

To see how the BSS model works, it is helpful to express the size of the PRP effect in terms of the durations of the component stages of Task 1 and Task 2. As shown in the Appendix:

$$\text{PRP effect} = 1A + 1B - 2A - \text{SOA}_{\text{short}}.$$

From this PRP equation, it follows that PRP reduction is caused by Task 1 practice but not by Task 2 practice. The predicted effects of Task 1 practice are straightforward: Task 1 practice should reduce the duration of Stages 1A and 1B and therefore should also reduce the PRP effect. The predicted effects of Task 2 practice, however, are somewhat counterintuitive. The only term in the PRP equation influenced by Task 2 practice is the duration of Stage 2A,[3] and there is a negative sign before this term. Hence, if Task 2 practice shortens Stage 2A, it would actually increase the size of the PRP effect (although, as discussed later, practice probably has little effect on Stage 2A). In sum, Task 1 practice should reduce the PRP effect, whereas Task 2 practice should not reduce the PRP effect. This interesting prediction was the primary inspiration for the new transfer experiments reported later in this article.

A plausible and very tractable subcase of the BSS model is obtained by adding the more specific assumption that only the central stages (1B and 2B) of each task become shorter with practice. We will refer to this possibility as the *bottleneck model with central stage shortening* (BCSS). This added assumption, even if not exactly true, is likely to be a close approximation to the truth given the tasks used in our studies. The reasoning behind this assumption is that the central stages (e.g., response selection) involve a novel, unpracticed mapping of stimuli onto responses, whereas the input stages (e.g., character identification) and output stages (e.g., speaking, button pressing) are already highly familiar. It stands to reason that the stages most sensitive to practice will be those that were not well practiced to begin with (i.e., the central stages). For supporting evidence, see Fletcher and Rabbitt (1978), Mowbray and Rhoades (1959), Pashler and Baylis (1991), and Welford (1976).

Given the assumption that practice has no effect on the noncentral stages (1A, 2A, 1C, 2C), the decrease in the PRP effect with practice should be due entirely to the decrease in the duration of Stage 1B (see the previous PRP equation). At the same time, the decrease in RT1 with practice would also be due entirely to the decrease in the duration of Stage 1B. Hence, the BCSS model predicts that the PRP effect and RT1 should drop by roughly the same amount across sessions. The VRJ data confirmed this prediction with a surprising degree of precision.

The BSS model and the more specific BCSS model are attractive not only because they account nicely for the VRJ data but also because they are simple and have high a priori plausibility. The large declines in RT1 and RT2 with practice (approximately 300 ms) could not have occurred unless at least some of the component stages shortened dramatically. Furthermore, the observed shortening of Task 1 stages with practice would appear, ceteris paribus, to necessarily result in a very large reduction in the PRP effect.

On the other hand, there was tentative evidence that stage shortening was not the only effect of practice in the VRJ study. Following extensive practice, Task 2 S-R compatibility (a variable assumed to influence the response selection stage) produced somewhat smaller effects at short SOAs than at long SOAs. Although this interaction did not reach statistical significance, it provides a tantalizing hint that compatibility effects were absorbed into cognitive slack. Hence, it is possible that some of the Task 2 response selection stage was automatized with practice (i.e., was carried out in parallel with central operations on Task 1).

Study Overview

The purpose of this study was to learn more about when and why practice reduces the PRP effect. We conducted three new experiments on the VRJ participants, taking advantage of the enormous amount of training these individuals had received (over 14,000 trials each). We pursued three specific goals: (a) to further test the BSS model, which predicts that PRP reduction is due primarily to practice on Task 1 not Task 2, (b) to see if some of the Task 2 response selection stage had become automatized with practice, and (c) to determine if differences in response modalities can explain why the VRJ study found much more PRP reduction than did previous studies.

Bottleneck Model With Stage-Shortening (BSS)

One specific goal of the present study was to put the BSS model, which accounted well for the VRJ findings, to a stricter test. According to this model, practice reduces stage durations but does not eliminate the processing bottleneck. Thus, the bottleneck exists both before and after practice.

As discussed earlier, the BSS model asserts that PRP reduction is due to practice on Task 1, not Task 2. We tested this claim in two new transfer experiments using 5 of the highly trained participants from the VRJ study. Each transfer experiment paired one of the highly practiced tasks from the VRJ study (either Task 1 or Task 2) with a new, unpracticed task. Experiment 1 paired the old, highly practiced Task 2 with a new Task 1. Because Task 1 was not highly practiced, this transfer experiment should show a large initial PRP effect. The PRP effect should then decline sharply with

---

[2] As discussed in VRJ, one of the 6 subjects (S.W.) showed little or no PRP effect after Session 12. S.W. might have learned to automatize the stages that formerly comprised the bottleneck (Stages 1B and 2B). Alternatively, it is possible that Stages 1B and 2B still existed and still could not be performed simultaneously, but Stage 1B was generally completed before Stage 2B was set to begin.

[3] We ignore, for the moment, the possibility that Task 2 practice might indirectly reduce Task 1 stage durations (e.g., by allowing participants to devote more pretrial preparation to Task 1).

further sessions of practice with the new Task 1. Experiment 2 paired the old, highly practiced Task 1 with a new Task 2. Because Task 1 was highly practiced, this transfer experiment should show a small PRP effect (roughly consistent with that observed at the end of the VRJ study) even in the very first session.

These transfer experiments also provide a further opportunity to evaluate the Task 1 carryover and Task 2 absorption predictions of the central bottleneck model and the BSS model. In addition, these new experiments will provide a further test of the more specific BCSS model in which practice is assumed to affect only the durations of central stages. As discussed previously, this model predicts that declines in the PRP effect with practice should equal the declines in RT1. In other words, a plot of the PRP effect versus RT1 across practice should have a slope of about 1.

In the General Discussion section, we contrast this model with a task-integration model, which says that practice leads to efficient integration of the two tasks. The central prediction of the task-integration model is that learning should transfer poorly to any new pair of tasks on which the participant has not been trained. We also consider a learned-automaticity model, which says that practice leads to complete task automatization (i.e., bypassing of the central bottleneck). The central prediction of this model is that learning should transfer well to new dual-task situations, provided that as least one of the two tasks has previously been automatized.

Automatization of Bottleneck Substages

A second goal of this study was to follow up on the hints in the VRJ data that Task 2 S-R compatibility effects were smaller at short SOAs than at long SOAs. This apparent absorption into cognitive slack would indicate that extensive practice allowed some of the Task 2 response selection stage to be carried out in parallel with the central stages of Task 1 (i.e., to be partially automatized). However, the effect did not reach significance in the VRJ study. This lack of significance might have occurred because the bottleneck delay had become very small, leaving insufficient cognitive slack to strongly absorb the S-R compatibility effects. As will be seen, Experiment 1 (new Task 1/old Task 2) produced a large bottleneck delay, leaving plenty of cognitive slack to absorb the effects of S-R compatibility. Thus, Experiment 1 should provide an excellent opportunity to see if the Task 2 response selection stage really had become partially automatized.

Role of Response Modalities

A third major goal of this study was to determine why VRJ found a much greater reduction in the PRP effect with practice than did previous studies. One likely explanation is that the VRJ tasks required one manual response and one vocal response, whereas previous investigators required two manual responses. As discussed earlier, there are reasons to believe that manual-manual designs result in response conflicts that persist with practice. To investigate this issue, Experiment 3 replicated the VRJ study by using essentially the same methods and the same tasks but with the requirement to make manual responses to both tasks. If the empirical discrepancy between the VRJ study and previous studies is due to differences in response modalities, then this manual-manual version of the VRJ experiment should show a relatively large initial PRP effect that declines only modestly with further practice.

Experiment 1:
Transfer to Design With New Task 1 / Old Task 2

Experiment 1 transferred 5 of the 6 highly trained participants from VRJ (S.W. was unavailable) to a design in which a new Task 1 was paired with the old Task 2 (see Table 1). In the VRJ study, participants determined whether a single tone was high or low in pitch on Task 1, whereas in Experiment 1, participants determined whether a pair of tones were same or different in pitch. Note that the input and output modalities of this new Task 1 are the same as those of the task it replaced (Task 2). Presumably, any increase in the size of the PRP effect would therefore be attributable to the novelty of Task 1 rather than to a change in the input or output modalities.

The BSS model asserts that it is practice on Task 1, not practice on Task 2, that causes PRP reduction. Because Task 1 was not highly practiced prior to this transfer experiment, the BSS model predicts a relatively large initial PRP effect. This effect should then decline relatively rapidly with further practice on the new Task 1 (i.e., as the Task 1 central stages become shorter).

Table 1
Task 1 and Task 2 Judgments Used in Van Selst, Ruthruff, and Johnston (1999; VRJ) and in Experiments 1–3

	Task 1		Task 2	
Experiment	Tone judgment	Response modality	Letter judgment	Response modality
Training (VRJ)	High/low	Vocal	ABCD1234	Manual
Experiment 1	Same/different*	Vocal	ABCD1234	Manual
Training refresher	High/low	Vocal	ABCD1234	Manual
Experiment 2	High/low	Vocal	XY*	Manual
Training refresher	High/low	Vocal	ABCD1234	Manual
Experiment 3	High/low	Manual*	ABCD1234	Manual

Note. The experiments are listed in chronological order, from oldest to newest. Asterisk (*) denotes a change in the design relative to the training design (VRJ).

To further test the BSS model, we also evaluated the Task 1 carryover and Task 2 absorption predictions. In addition, we tested the BCSS model, which includes the extra assumption that only the central stages become shorter with practice. This model predicts that PRP reduction should closely track RT1 reduction across sessions of practice. Finally, we looked for absorption of Task 2 S-R compatibility effects at short SOAs, which would indicate that participants had automatized at least some of the Task 2 response selection stage.

Method

Five participants performed four sessions of a PRP experiment (generally one session per day, excluding weekends). These same individuals previously performed 36 sessions of a dual-task experiment reported by VRJ, which was nearly identical to Experiment 1 except that it involved a different Task 1 judgment (see Table 1). One additional difference is that the SOAs in Experiment 1 {−33, 17, 200, 800} were a subset of those used in the VRJ study {−33, 17, 100, 200, 400, 800}. The purpose of this change was to increase the amount of data obtained at the longest and shortest SOAs, which are used to measure the size of the PRP effect.

Participants. Participants E.R. and M.V. are coauthors of this article.[4] The remaining participants (V.L., J.C., M.R.) were recruited from work-study programs at the National Aeronautics and Space Administration, Ames Research Center (Moffett Field, CA).

Stimuli. The stimulus for Task 1 was a pair of tones presented for 84 ms each, with an interstimulus interval of 150 ms. The frequency (Hz) of the first tone was selected at random from the set {1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600}. On half of the trials, the second tone was identical in pitch to the first tone (*same* trials). On the other half of the trials, the frequency of the second tone was either 0.2, 0.6, 1.4, or 1.8 times the frequency of the first tone (*different* trials). RT1 and the SOA were measured relative to the onset of the second tone. The stimulus for Task 2 was a single alphanumeric character drawn at random from the set {1, 2, 3, 4, A, B, C, D}. The characters were presented in Times Roman font at a viewing distance of about 65 cm. All characters fit within a rectangular area of 1.41° × 0.94° of visual angle. The background was white, and the characters were black (high-contrast condition) or gray (low-contrast condition).

Apparatus. Stimulus presentation and timing were performed by a Compaq 386 microcomputer equipped with a Votan voice recognition system and a Schmitt trigger voice key.

Procedure. Participants responded to the pair of tones with a vocal response ("same" or "differ") and responded to the identity of the alphanumeric character by pressing the *H, J, K,* or *L* key on a standard keyboard, using the fingers of their right hands. For some of the participants (M.V., J.C., M.R.), the letters A, B, C, and D were mapped in alphabetic order onto the four response keys from left to right (i.e., compatibly). The numbers, meanwhile, were mapped in an incompatible, nonsequential order (3, 1, 4, 2) onto the same four response keys. For the remaining participants (E.R., V.L.), numbers were mapped compatibly (1, 2, 3, 4) but letters were mapped incompatibly (C, A, D, B). Instructions emphasized the importance of responding quickly and accurately to both tasks. Particular emphasis was placed on the speed of Task 1 responses.

Each trial began with the presentation of a fixation cross for 500 ms, followed by a random foreperiod (100, 150, 200, or 250 ms). The SOA between the second of the two tones and the alphanumeric character, S2, was −33, 17, 200, or 800 ms. S2 remained present until the participant responded to it. If either response was incorrect, a message indicated the task on which the error had been made. Also, if the participant responded within 100 ms of stimulus onset, a "too early" message was displayed. If the participant failed to respond within 2,500 ms of stimulus onset, a "too slow" message was displayed. The intertrial interval was 750 ms.

Each session consisted of 35 warm-up trials followed by 525 experimental trials. The session was broken into seven blocks of 80 trials each, separated by short breaks. During each break, the computer provided feedback on the average speed of Task 1 and the accuracy of both Task 1 and Task 2.

Analyses. We conducted separate analyses of variance (ANOVAs) on mean RT1, RT2, Task 1 error rate, and Task 2 error rate. Before conducting RT analyses, we first eliminated trials containing an error on either task. We then removed all trials with an RT of less than 100 ms, followed by any trial in which either RT1 or RT2 was identified as an outlier (less than 5.6% of trials in each experiment) using a modified recursive outlier elimination procedure with moving criterion (Van Selst & Jolicoeur, 1994). All analyses used an alpha level of .05.

An additional set of paired t tests were conducted to compare the size of the PRP effects observed in the VRJ study and in Experiments 1–3 of this article. Although the present experiments included both a −33-ms SOA and a 17-ms SOA, the early phases of the VRJ study included only the 17-ms SOA. Therefore, to facilitate comparisons with VRJ, the present PRP effects were always measured using the 17-ms SOA as a baseline rather than the −33-ms SOA. Furthermore, we report only VRJ data from the 5 individuals who also participated in the present experiments. Thus, all comparisons between experiments are based on the same participants and the same SOAs.

Results and Discussion

Figure 2 shows the mean PRP effect (measured as $\mathrm{RT2}_{\mathrm{SOA}\ 17} - \mathrm{RT2}_{\mathrm{SOA}\ 800}$), RT1, and RT2 as a function of practice session. For purposes of comparison, we also show the initial and final values from the VRJ study (dotted lines in Figure 2). The RT2 data in Figure 2 come from the long SOA condition only, to show the improvement in baseline Task 2 performance with practice (i.e., apart from any reductions in dual-task interference). Figure 3 shows mean RT1, RT2, and several factor effect sizes as a function of SOA, pooled across all four sessions. The top panel shows RT1 (dashed line) and RT2 (solid line); the second panel shows the effect of a Task 1 factor (*same* vs. *different* trials) on RT1 (dashed line) and on RT2 (solid line); the third panel shows the effect of Task 2 contrast on RT2; the fourth panel shows the effect of Task 2 S-R compatibility on RT2. Table 2 shows the initial PRP effect (i.e., Session 1) for each participant in Experiment 1, along with the same data from VRJ and Experiments 2 and 3.

Before the size of the PRP effect is discussed, it is important to determine whether Task 2 performance was adversely affected by the introduction of a new Task 1. Mean RT2 at the longest SOA was 474 ms in Session 1 of Experiment 1, compared to 457 ms in the final phase of the VRJ experiment (see Figure 2). This difference was not significant, $F(1, 4) < 1$. It appears, therefore, that Task 2 learning transferred well despite being paired with a new Task 1.

Initial PRP effect. The initial PRP effect was relatively large (194 ms). This effect is much greater than the 50-ms PRP effect these participants produced at the end of the VRJ experiment, $t(4) = 9.4, p < .001$. In fact, the PRP effect for each participant in Experiment 1 was at least 100 ms greater than it was at the end of the VRJ experiment. As predicted by the BSS model, prior practice on Task 2 alone was not sufficient to dramatically reduce the PRP effect.

---

[4] The data produced by M.V. and E.R. were not qualitatively different from those produced by the other participants (see, e.g., Table 2).

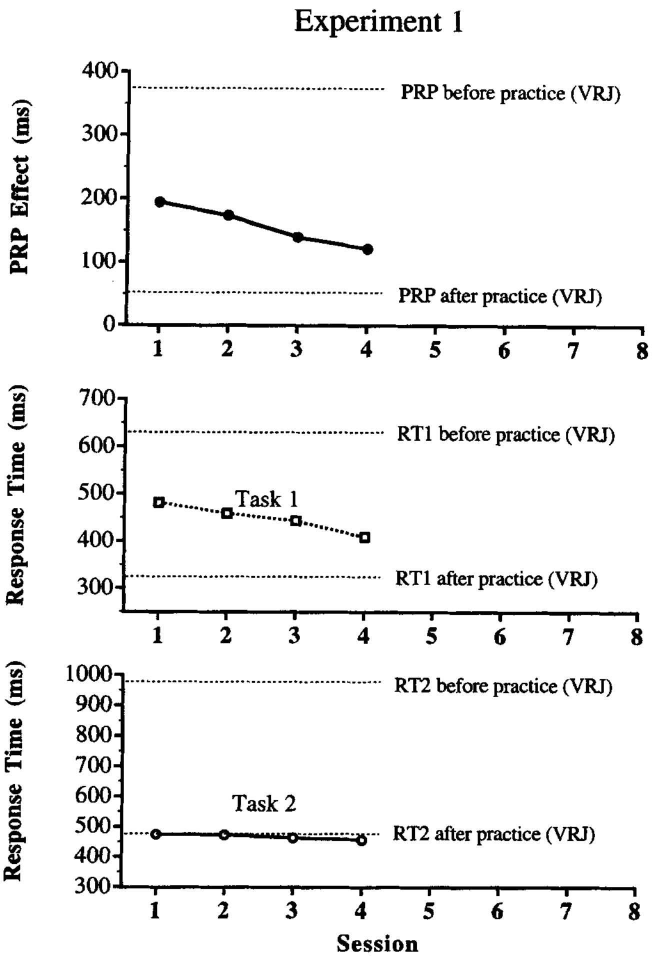

Figure 2. The size of the psychological refractory period (PRP) effect, mean Task 1 response time (RT1), and mean Task 2 response time (RT2; at the longest stimulus onset asynchrony) in Experiment 1 as a function of session. The dashed lines show the mean effect sizes before and after practice in the Van Selst et al. (1999; VRJ) experiment.

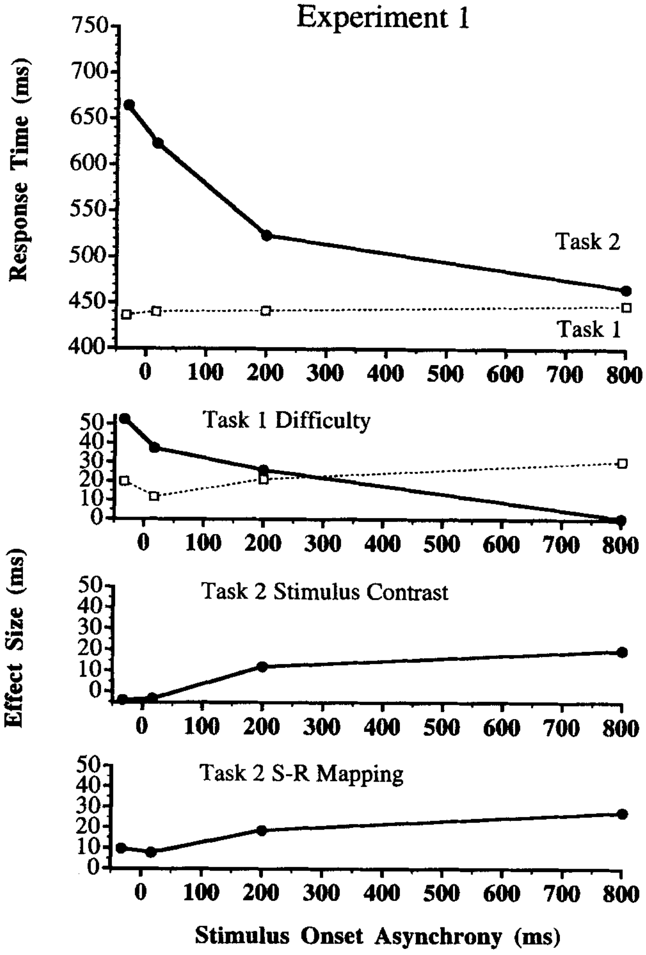

Figure 3. Performance measures in Experiment 1 averaged across all four sessions as a function of stimulus onset asynchrony (SOA). The upper panel shows the effect of SOA on Task 1 response times (RT1, dashed line) and Task 2 response times (RT2, solid line). The second panel shows the effect of Task 1 difficulty on RT1 (dashed line) and RT2 (solid line). The third panel shows the effect of Task 2 stimulus contrast on RT2. The bottom panel shows the effect of Task 2 stimulus–response (S-R) mapping difficulty on RT2.

Although learning clearly did not transfer completely from the VRJ study, there might have been partial transfer of learning. Indeed, the initial PRP effect in Experiment 1 was 179 ms less than that found in the initial session of the VRJ study (see Figure 2). This observation must be qualified, however, by noting that the initial mean RT1 in Experiment 1 was 151 ms less than that found in the VRJ study (477 vs. 628 ms). As noted previously, bottleneck models predict that reductions in RT1 will lead to reductions in the PRP effect. This reason alone is sufficient to account for the modest PRP reduction observed in Experiment 1 relative to the initial sessions in the VRJ study. Put another way, there is no compelling evidence that prior Task 2 practice had any direct influence on the size of the PRP effect (e.g., by bypassing the bottleneck altogether).

Why was Task 1 performed so quickly even though it was unpracticed? At least some of the RT1 speedup can be attributed to a speed–accuracy trade-off on Task 1 (see the *Task 1 response time* section below). Furthermore, this new Task 1 might have been easier, participants might have benefited from general practice with the experimental apparatus and procedures, or both. It is also conceivable that well-learned Task 2 judgments require little pretrial preparation, allowing participants to instead devote their preparation to Task 1. Thus, Task 1 would have been better prepared in

Table 2

Initial Psychological Refractory Period Effects for Each Experiment by Participant (Measured Relative to the 16-Millisecond Stimulus Onset Asynchrony)

Participant	VRJ	Experiment 1	Experiment 2	Experiment 3
M.V.	416	245	112	418
M.R.	517	200	106	271
J.C.	278	258	146	463
E.R.	336	139	77	243
V.L.	317	124	48	399

Note. VRJ = Van Selst, Ruthruff, and Johnston, 1999.

Experiment 1 than it was in the initial sessions of the VRJ study (where Task 2 was still unpracticed), resulting in a faster mean RT1 (see Gottsdanker, 1980). According to this hypothesis, Task 2 practice did serve to reduce the PRP effect somewhat, albeit indirectly. Note that such an effect is in no way contrary to the spirit of the bottleneck model. Further research is needed to determine just how much of the partial transfer of learning is due (directly or indirectly) to prior Task 2 practice and how much is due to extraneous factors such as general practice effects.

The effect of further practice. As predicted by the BSS model, the PRP effect decreased sharply with further dual-task practice (see Figure 2), $F(9, 36) = 5.19, p < .001$. By the fourth session, the PRP effect had dropped from 194 ms down to only 121 ms. Note that Task 2 performance had approached asymptote long before Experiment 1 began, as evidenced by the fact that RT2 decreased only very slightly across sessions. Hence, the sharp decrease in the PRP effect across sessions can be attributed to further practice on the new Task 1.

PRP effect versus RT1. The BSS model, coupled with the assumption that practice shortens only the central stages (i.e., the BCSS model), predicts that a plot of the PRP effect versus RT1 across sessions should have a slope of about 1. Figure 4 shows the observed PRP effect as a function of mean RT1 across the four sessions of Experiment 1 (square symbols). Also shown are data from the initial 18 sessions of the VRJ study (plus symbols). The best linear fits of the data do indeed show slopes close to 1.0 both for Experiment 1 (slope = 1.03) and for the VRJ experiment (slope = 1.01). Thus, these data provide strong support for the BCSS model.

Note that the linear fits to the VRJ study (Sessions 1–18) and Experiment 1 appear to have different intercepts. Experiment 1 data show about 50 ms less PRP effect for any given RT1. According to the BSS model, there are two obvious explanations for this intercept difference. One possibility is that the duration of the Task 1 output stage (1C) was greater in Experiment 1 (same vs. different) than in VRJ (high vs. low), leading to an increase in RT1 with no increase in the PRP effect. Alternatively, the duration of Stage 2A might have been longer in Experiment 1 than in the VRJ study, leading to a decrease in the PRP with no decrease in RT1. At first blush, the proposed increase in Stage 2A seems unlikely given that participants in Experiment 1 had more practice with this Task 2 than they did in VRJ's experiment. On the other hand, it is possible that part of the Task 2 response selection stage had become automatized sometime between VRJ's experiment (Sessions 1–18) and Experiment 1; participants might have learned to perform part of the Task 2 response selection stage during the bottleneck delay (see *Locus of the bottleneck* section below for supporting evidence). In effect, partial automatization would subtract from the duration of the bottleneck stage (2B) while adding to the duration of the prebottleneck stage (2A).[5]

Task 1 carryover prediction. The central bottleneck model predicts that Task 1 difficulty effects should carry over onto RT2 at short SOAs but not at long SOAs. Consistent with this prediction, RT1 was 21 ms slower on *different* trials than on *same* trials, and this effect carried over onto RT2 more at short SOAs (53 ms) than at long SOAs (0.4 ms), $F(3, 12) = 6.19, p < .01$ (see Figure 3). Note that the measured amount of carryover onto RT2 at the short SOA (53 ms) exceeds the size of the effect on RT1 (21 ms). This unanticipated result is likely due to an underestimation of the true Task 1 difficulty effect on RT1. Tests of our equipment show that the voice key requires more time (~25 ms) to detect the onset of the word *same* than to detect the onset of the word *differ.*

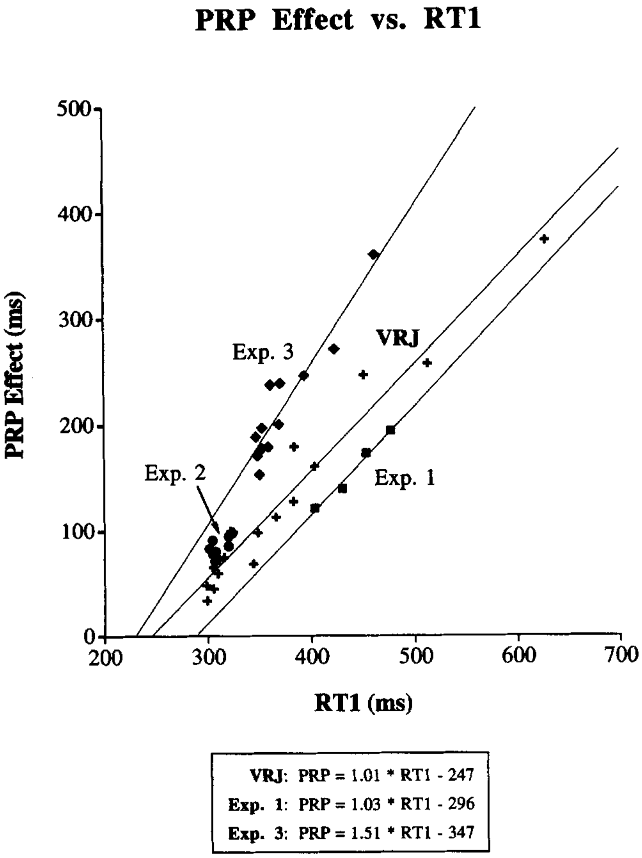

Figure 4. Size of the psychological refractory period (PRP) effect as a function of mean Task 1 response time (RT1) across sessions in Van Selst et al. (1999; VRJ; plus signs), Experiment 1 (squares), Experiment 2 (circles), and Experiment 3 (diamonds). Each data point represents the group average for one session. The solid lines through each data set show the best linear fit. The linear fit for the Experiment 2 data is not shown due to range restriction. Exp. = Experiment.

Task 2 absorption prediction. The central bottleneck model predicts that manipulations of prebottleneck stages of Task 2 should have a much smaller effect on RT2 at short SOAs than at long SOAs (i.e., the factor effects should be absorbed into cognitive slack). Consistent with this prediction, the effect of Task 2 stimulus contrast was 20 ms at the longest SOA but was only 4 ms at the shortest SOA (see Figure 3, third panel). This trend was only marginally significant in the overall analysis (i.e., using all four SOAs), $F(3, 12) = 2.87, p = .08$, but was statistically significant in a contrast analysis including only the shortest and longest SOAs, $F(1, 4) = 15.7, p < .05$.

Locus of the bottleneck. If the Task 2 response selection stage had become automatized, then Task 2 S-R compatibility effects

---

[5] The automated part of the response selection stage is now, by definition, "prebottleneck." Thus, it adds to the duration of the prebottleneck stage (Stage 2A).

should have been smaller at short SOAs due to absorption into slack. As in the VRJ data, we did observe a trend toward smaller compatibility effects on RT2 at short SOAs: The compatibility effect was 27 ms at the longest SOA but was only 10 ms at the shortest SOA. Although the underadditive interaction was not statistically significant in VRJ's study, it did reach significance in Experiment 1, $F(3, 12) = 3.68$, $p < .05$. These findings suggest that at least part of the Task 2 response selection stage had become automatized with practice (i.e., took place while Task 1 central operations were still underway).[6]

Task 2 error rates. Task 2 error rates were too low (2.1%) to permit meaningful analysis.

Task 1 response time. RT1 decreased with practice, $F(3, 12) = 6.10$, $p < .01$. No other RT1 effects were statistically significant. Participants reported that they felt some urgency to respond to the new Task 1 as quickly as they had been responding to the old Task 1 in the VRJ study, even though this new Task 1 was unpracticed and therefore much harder. Consistent with these self-reports, we found evidence of a speed–accuracy trade-off between experiments: Participants responded more rapidly in the four sessions of Experiment 1 than they did in the first four sessions of the VRJ study (441 ms vs. 499 ms), but they made far more errors (12.2% vs. 5.2%).

Task 1 error rates. Participants made an average of 12.2% errors on Task 1. Task 1 error rates decreased with practice, $F(3, 12) = 4.3$, $p < .05$. Error rates also decreased with increasing SOA, $F(3, 12) = 8.3$, $p < .01$: Participants made 13.9% errors at the shortest SOA and 10.7% errors at the longest SOA. The RT1 data show a nonsignificant trend toward faster responses at short SOAs (436 ms) compared to long SOAs (448 ms), $F(3, 12) = 1.58$, $p > .2$, so the difference in error rates across SOAs might be due to a speed–accuracy trade-off. It is conceivable that, at short SOAs, the appearance of S2 induces participants to complete Task 1 processing more quickly (e.g., to free up resources needed to process S2), at the expense of making more errors.

Experiment 2:
Transfer to Design With Old Task 1 / New Task 2

Experiment 2 transferred the 5 participants from the VRJ study to a design in which the old Task 1 (used by VRJ) was paired with a new Task 2 (see Table 1). The new Task 2 was to discriminate between the letters *X* and *Y* and to respond by manually pressing one of two response keys. Note that this task used the same input and output modalities as the VRJ task that it replaced.

The BSS model predicts that it is prior practice on Task 1, not prior practice on Task 2, that causes PRP reduction. Because the Task 1 used in this transfer experiment was already highly practiced, we should observe a small PRP effect, even in the very first session. Furthermore, the PRP effect should decline very little across subsequent sessions of practice in this design (except to the degree that there is some further stage shortening on Task 1).

Note that the goal of Experiment 2 was to measure transfer of learning from the VRJ study to a design with a new Task 2; at this point, however, participants had just completed four sessions of the Experiment 1 design. Therefore, to reinstate the learning from the VRJ study, we retrained participants on the tasks from that study for two full sessions prior to beginning the eight transfer sessions of Experiment 2.

Method

Except where noted, the method was identical to that of Experiment 1. See Table 1 for a brief summary of the relations between Experiments 1–3 and the VRJ study.

Stimuli. As in the VRJ study, the stimulus for Task 1 was one of four tones presented for a duration of 150 ms. The two tones highest in pitch (3125 and 625 Hz) were classified as *high* tones, and the two tones lowest in pitch (400 and 80 Hz) were classified as *low* tones. The stimulus for Task 2 was the letter *X* or the letter *Y*. Stimulus discriminability was varied by altering the fonts of the letters on half of the trials (the vertex of the *Y* was lowered and the vertex of the *X* was raised). This factor had negligible effects ($M < 2$ ms), however, and hence was omitted from the final analyses.

Procedure. Participants responded to Task 1 by saying "high" in response to the two high-pitched tones and "low" in response to the two low-pitched tones. Participants responded to Task 2 by pressing the *1* key for the letter *X* or by pressing the *2* key for the letter *Y*, using the numeric keypad.

Results and Discussion

Figure 5 shows the mean PRP effect, RT1, and RT2 across the eight experimental sessions. For purposes of comparison, we also show the initial and final values from the VRJ study (dotted lines). The RT2 data in Figure 5 come from the long SOA condition only to show the improvement in baseline Task 2 performance with practice (i.e., apart from any reductions in dual-task interference). Figure 6 shows mean RT1, mean RT2, and the Task 1 difficulty effect as a function of SOA, pooled across all eight sessions. The top panel shows RT1 (dashed line) and RT2 (solid line), and the second panel shows the effect of Task 1 difficulty (extreme vs. intermediate tone frequencies).

Before the size of the PRP effect is discussed, it is important to examine whether Task 1 performance was adversely affected by the introduction of a new Task 2. Mean RT1 was 322 ms in Session 1 of Experiment 2, compared to 318 ms in the final phase of VRJ's study. This difference was not significant, $F(1, 4) < 1$. Therefore, Task 1 learning transferred well despite being paired with an unfamiliar Task 2.

Initial PRP effect. As predicted by the BSS model, the initial PRP effect in Experiment 2 was small (98 ms). This PRP effect is much smaller than the initial PRP effect of 373 ms found in the VRJ study, $t(4) = 6.1$, $p < .01$.

Note that the present Task 2 was fairly easy. It is conceivable that it was, in fact, too easy to produce a large PRP effect. This possibility seems unlikely, given that PRP effects generally do not

---

[6] It is possible that the observed underadditivity was due not to automatization of the response selection stage, but rather to small (unintended) effects of S-R compatibility on the character identification stage. For instance, suppose that participants who made incompatible responses to letter stimuli identified the letter stimuli more carefully (i.e., more slowly) than the number stimuli, perhaps to compensate for the greater likelihood of response selection errors. Because character identification presumably precedes the bottleneck (e.g., Van Selst & Johnston, 1997), this portion of the compatibility effect (say, 10–20 ms) could be absorbed into cognitive slack. Early in practice the portion subject to absorption would constitute a negligible fraction of the total compatibility effect (about 230 ms), but after practice this portion might very well constitute a large proportion of the total compatibility effect (only 23 ms).

depend much on Task 2 difficulty[7]; in fact, PRP effects are usually larger when the prebottleneck stages of Task 2 are especially short (see McCann & Johnston, 1992; Pashler & Johnston, 1989). Nevertheless, to address this concern, we ran 22 naïve control participants for one session each in the design of Experiment 2. These participants, who had no previous experience with Task 1 or Task 2, produced a very large mean PRP effect (470 ms). Even the smallest PRP effect of the 22 control participants was 291 ms, whereas the largest PRP effect of the 5 highly practiced participants in Experiment 2 was only 146 ms. Thus, we conclude that the small PRP effect found in Experiment 2 was due to prior practice on Task 1 rather than to the ease of the new Task 2.

The initial PRP effect of 98 ms found in Experiment 2 was significantly less than the initial PRP effect of 194 ms found in Experiment 1, $t(4) = 7.6$, $p < .01$. In fact, the PRP effect for each individual was at least 62 ms smaller than it was in Experiment 1 (see Table 2). This comparison suggests that prior practice on Task 1 was more helpful in reducing the PRP effect than was prior practice on Task 2, just as predicted by the BSS model.[8]

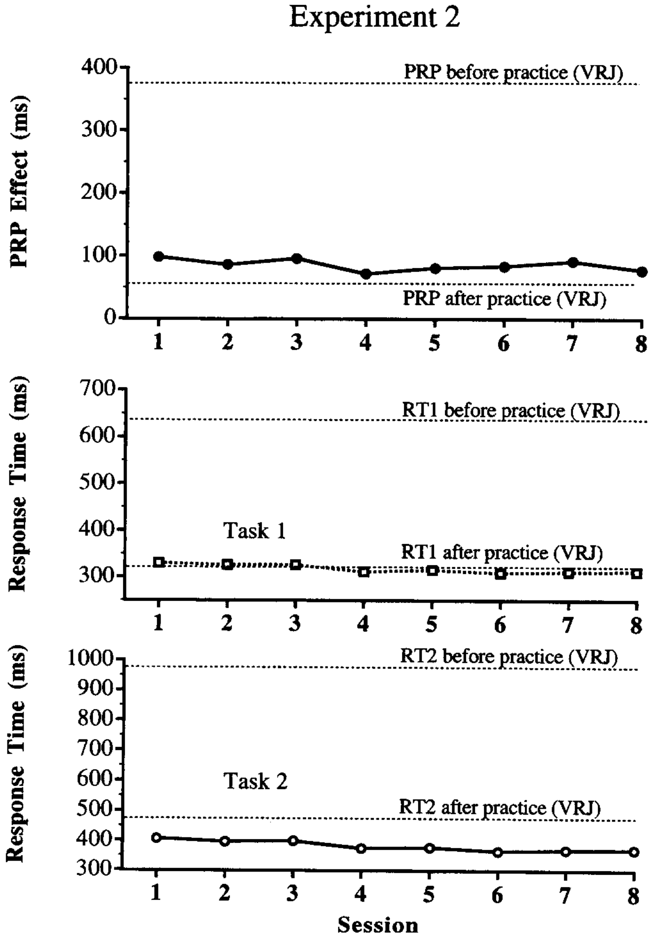

Figure 5. The size of the psychological refractory period (PRP) effect, mean Task 1 response time (RT1), and mean Task 2 response time (RT2; at the longest stimulus onset asynchrony) in Experiment 2 as a function of session. The dashed lines show the mean effect sizes before and after practice in the Van Selst et al. (1999; VRJ) experiment.

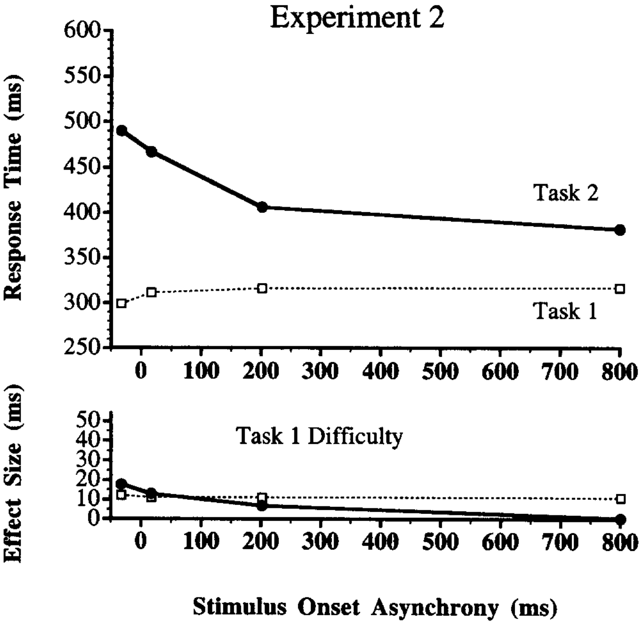

Figure 6. Performance measures in Experiment 2 averaged across all eight sessions as a function of stimulus onset asynchrony (SOA). The upper panel shows the effect of SOA on Task 1 response times (RT1; dashed line) and Task 2 response times (RT2; solid line). The second panel shows the effect of Task 1 difficulty (intermediate vs. extreme tone frequencies) on RT1 (dashed line) and RT2 (solid line).

Although prior practice on Task 1 clearly helped to drastically reduce the PRP effect, the initial PRP effect (98 ms) was not quite as small as the 50-ms effect these participants produced after practice in the VRJ study, $t(4) = 8.3$, $p < .01$. This 48-ms elevation in the PRP effect might have occurred because the present Task 2 involved only two stimuli (X, Y), whereas Task 2 of VRJ involved eight different stimuli (A, B, C, D, 1, 2, 3, 4). The smaller number of stimuli might have caused a reduction in the duration of Stage 2A, which in turn would have elevated the size of the PRP effect relative to the effect in the VRJ study (see the PRP equation displayed previously). This hypothesis is supported by the finding that the control participants produced an especially large PRP effect.

Effect of further practice. Because Task 1 was highly practiced prior to Experiment 2, the BSS model predicts that the PRP effect should not decline much after the initial session (except to the degree that there is some further decrease in RT1). Consistent with this prediction, the PRP effect declined very little across the next seven sessions (see Figure 5). Numerically, the PRP effect did

---

[7] An exception to this rule, noted earlier, can occur if Task 2 is ideomotor compatible. However, the present Task 2 is clearly not ideomotor compatible.

[8] This comparison is confounded somewhat because participants had more dual-task practice prior to Experiment 2 than they did prior to Experiment 1. Note, however, that participants were very highly practiced even before Experiment 1 began. Therefore, the small increment in the amount of practice from Experiment 1 to Experiment 2 was probably inconsequential.

show a nonsignificant decline from 98 to 78 ms, $F(21, 84) < 1$, but this decline could easily be attributed to a similar decline in RT1 from 322 to 305 ms.

PRP effect versus RT1. Figure 4 shows the observed PRP effect as a function of mean RT1 across the eight sessions of Experiment 2 (circles). Because there was so little variation in the PRP effect and RT1 across sessions, we did not attempt to find the best linear fit to the data. Note that the Experiment 2 data appear to lie slightly above those from the VRJ study, although the data sets do overlap to some degree. If real, this difference might reflect an especially short Stage 2A in Experiment 2. This shortening of Stage 2A would have increased the PRP effect while having no impact on RT1.

Task 1 carryover prediction. Task 1 responses to the extreme tones (80 and 3125 Hz) were 11 ms faster than responses to the intermediate tones (400 and 625 Hz), $F(1, 4) = 9.1$, $p < .05$. As predicted by bottleneck models, we observed more Task 1 carryover onto RT2 at the shortest SOA (18 ms) than at the longest SOA (0.6 ms), $F(3, 12) = 4.19$, $p < .05$ (see Figure 6).

Task 2 error rates. Participants made an average of 5.3% errors on Task 2. Task 2 error rates did not vary significantly with any of our experimental factors.

Task 1 response time. Mean RT1 depended on SOA, $F(3, 12) = 47.4$, $p < .001$; mean RT1 was 299 ms at the shortest SOA and 317 ms at the longest SOA. Further, the effects of session on RT1 interacted with the effects of Task 2 S-R mapping, $F(7, 28) = 3.1$, $p < .05$. In addition, there was a three-way interaction on RT1 between session, SOA, and Task 1 difficulty, $F(21, 84) = 2.0$, $p < .05$. Because the latter two effects were not monotonic across SOA and were not significant in VRJ's study or in the other experiments reported in this article, we do not attempt to interpret them.

Task 1 error rates. Participants made an average of 3.7% errors on Task 1. The only variable to have a significant effect on Task 1 error rates was Task 1 difficulty, $F(7, 28) = 2.81$, $p < .05$; participants made 1.9% errors to the extreme tone frequencies and 5.5% errors to the intermediate tone frequencies.

Discussion of Experiments 1 and 2

The results of Experiments 1 and 2 clearly show that prior practice on Task 1 is more beneficial than prior practice on Task 2. When Task 2 was highly practiced and Task 1 was not (Experiment 1), participants produced relatively large PRP effects. When Task 1 was highly practiced and Task 2 was not (Experiment 2), participants produced small PRP effects. These results support the BSS model as an account of the effects of practice on dual-task interference. This model was further supported by confirmation of two key bottleneck model predictions: (a) Task 1 difficulty effects carried over fully onto RT2, and (b) Task 2 contrast effects appeared to be completely absorbed at short SOAs. Furthermore, declines in the PRP effect closely tracked declines in RT1, as predicted by the BCSS (a subcase of the BSS model in which practice shortens only the central stages). In the General Discussion section, we consider the possibility that practice also caused some of the Task 2 bottleneck substages to become automatized.

Experiment 3: Transfer to Design With Manual Responses to Both Tasks

The purpose of Experiment 3 was to determine why practice dramatically reduced the PRP effect in VRJ's study but had relatively little effect in previous studies. Although there are several differences in experimental design, perhaps the most salient difference is that VRJ used one vocal response task and one manual response task, whereas previous investigators used two manual response tasks. VRJ speculated that the use of manual responses on both tasks leads to response conflicts, which are particularly resistant to practice. To verify that the difference in Task 1 response modality was responsible for the discrepancy in results, we replicated the original VRJ design with only one change: Participants made manual rather than vocal responses to the Task 1 stimuli. Hence, both tasks required manual responses, as in previous PRP studies involving extended practice.

Task 2 was identical to Task 2 in VRJ's study. Task 1 was also identical to that of VRJ, except that rather than saying "high" and "low" in response to high- and low-pitched tones on Task 1, participants pressed one of two buttons labeled *high* and *low* (see Table 1). To ensure that the responses were compatible with the stimulus categories, the response button assigned to high-pitched tones (the *S* key) was located higher on the keyboard than the button assigned to low-pitched tones (the *X* key).

The 5 highly trained participants first completed two refresher sessions with the original VRJ tasks. They then completed 12 sessions of the new manual-manual version of the VRJ experiment. If the use of manual responses results in stubborn response conflicts, then these participants should produce relatively large PRP effects despite the very large amount of practice they received in VRJ's study on nearly the same pair of tasks (only the response to Task 1 was different) and despite 12 sessions of additional practice in this transfer experiment.

Method

Except where noted, the method was identical to that in Experiments 1 and 2. Task 1 was the high and low tone frequency judgment used in Experiment 2 and in the VRJ study, except that participants responded manually (*S* key for high-pitched tones, *X* key for low-pitched tones) instead of vocally. Task 2 was the same letter and number task (A, B, C, D, 1, 2, 3, 4) used in Experiment 1 and in the VRJ study (see Table 1). We used the full range of SOAs from Phase 3 of VRJ's study (−33, 17, 100, 200, 400, and 800 ms).

Results and Discussion

Figure 7 shows the mean PRP effect, RT1, and RT2 across the 12 experimental sessions. The RT2 data in Figure 7 come from the long SOA condition only to show the improvement in baseline Task 2 performance with practice (i.e., apart from any reductions in dual-task interference). Also shown, for the purpose of comparison, are the PRP data from the first 12 sessions of the VRJ study. Figure 8 shows mean RT1, mean RT2, and several factor effect sizes as a function of SOA, pooled across all 12 sessions. The top panel shows RT1 (dashed line) and RT2 (solid line); the second panel shows the effect of Task 1 difficulty (extreme vs. intermediate tone frequencies); the third panel shows the effect of Task 2

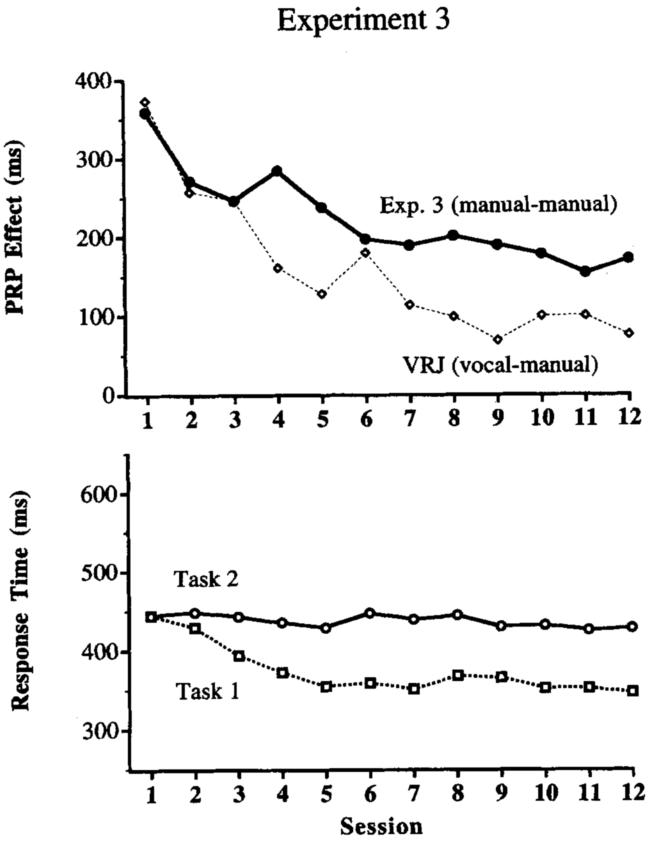

Figure 7. The size of the psychological refractory period (PRP) effect, mean Task 1 response time, and mean Task 2 response time (at the longest stimulus onset asynchrony) in Experiment 3 as a function of session. Also shown, for the purpose of comparison, is the mean PRP effect in the first 12 sessions of the Van Selst et al. (1999; VRJ) study (dotted line). Exp. = Experiment.

contrast; and the fourth panel shows the effect of Task 2 S-R compatibility.

Before the size of the PRP effect is discussed, it is important to examine whether Task 2 performance was adversely affected by the change in Task 1 (i.e., replacing the vocal response with a manual response). The mean long SOA RT2 was 444 ms in Session 1 of Experiment 3 compared to 457 ms in the final phase of VRJ's experiment. If anything, RT2 was slightly faster in Experiment 3, although this difference was not significant, $F(1, 4) < 1$. Thus, Task 2 learning transferred very well despite the change in the Task 1 response modality.

Initial PRP effect. The initial PRP effect was 359 ms, which is much greater than the 50-ms PRP effect these participants produced after practice in the VRJ study, $t(4) = 8.6, p < .001$. In fact, the PRP effect was at least 220 ms greater for each participant in Experiment 3 than it was after practice in VRJ's study. Note that the initial mean RT1 in Experiment 3 was 462 ms, whereas the mean RT1 at the end of the VRJ experiment was only 318 ms. According to bottleneck models, this 144-ms elevation in RT1 can account for roughly 144 ms of elevation in the PRP effect (i.e., Task 1 carryover). Therefore, the increase in RT1 can account for only about half of the 288 ms increase in the PRP effect observed in Experiment 3 relative to the VRJ study. In the General Discussion section, we consider several alternative explanations for this effect.

The initial PRP effect (359 ms) in Experiment 3 was, in fact, virtually identical to that found in the very first session of VRJ's study (373 ms), $t(4) = 0.19, p > .2$, even though participants had enormous amounts of prior practice with essentially these same tasks (only the Task 1 response modality was different) and with the general dual-task paradigm. This comparison makes it clear that learning from the VRJ study transferred remarkably poorly to this manual-manual design.

Effect of further practice. The PRP effect dropped with further practice on the new Task 1 S-R mapping, from 359 ms in Session 1 down to 171 ms in Session 12. This residual PRP effect, however, was still nearly 100 ms larger than that observed after the same amount of practice in the VRJ study (see Figure 7), $t(4) = 3.0, p <$

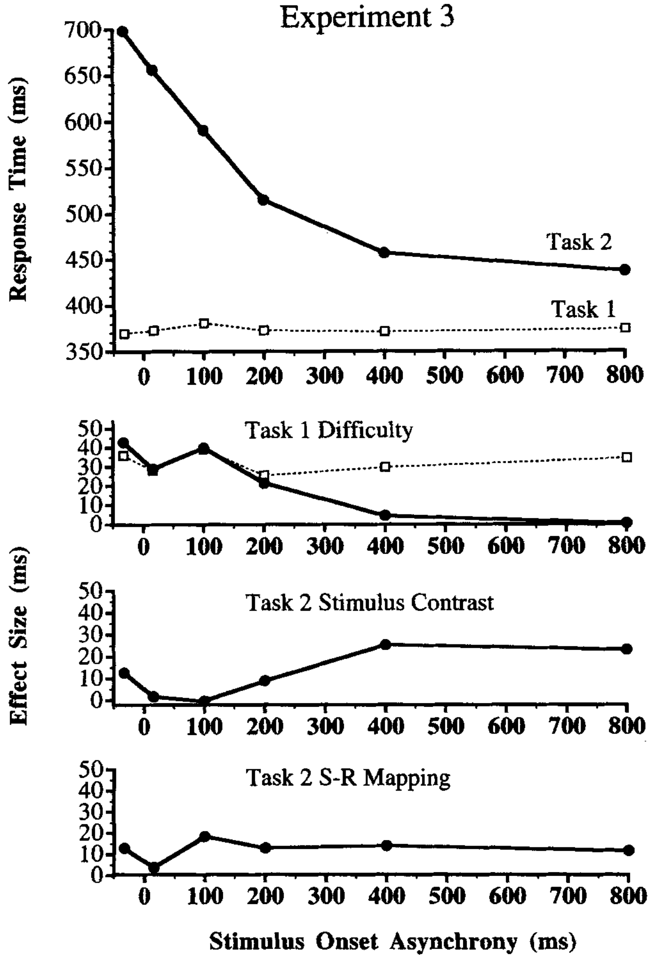

Figure 8. Performance measures in Experiment 3 averaged across all 12 sessions as a function of stimulus onset asynchrony (SOA). The upper panel shows the effect of SOA on Task 1 response times (RT1; dashed line) and Task 2 response times (RT2; solid line). The second panel shows the effect of Task 1 difficulty (intermediate vs. extreme tone frequencies) on RT1 (dashed line) and RT2 (solid line). The third panel shows the effect of Task 2 stimulus contrast on RT2. The bottom panel shows the effect of Task 2 stimulus–response (S-R) mapping difficulty on RT2.

.05. The amount of residual PRP interference is especially impressive given that even before Experiment 3 began these participants had received many thousands of trials of prior practice with the tone discrimination of Task 1—albeit using a different response modality. Thus, these results provide strong support for VRJ's speculation that the PRP reduction with practice is less dramatic when both tasks require manual responses.

PRP effect versus RT1. The BSS model, coupled with the assumption that practice shortens only the central stages (i.e., the BCSS model), predicts that a plot of the PRP effect versus RT1 across sessions should have a slope of about 1. Figure 4 shows the observed PRP effect as a function of mean RT1 across the 12 sessions of Experiment 3 (represented by diamonds). The best linear fit of these data has a slope 1.51, which appears to deviate from the predicted slope of 1.0. A t test comparing the slopes against 1.0, however, failed to reveal a significant difference, $t(4) = 1.2, p > .2$. Also note that the Experiment 3 data lie well above those from the other experiments, indicating that a given RT1 tended to produce much more interference in this experiment than it did in VRJ's study. In the General Discussion section, we consider several specific accounts for why these data appear to differ from those of VRJ and Experiments 1 and 2.

Task 1 carryover prediction. Mean RT1 to the extreme tones (80 and 3125 Hz) was 32 ms faster than mean RT1 to the intermediate tones (450 and 600 Hz), $F(1, 4) = 9.6, p < .05$. As predicted by bottleneck models, this effect carried over onto RT2 more at the shortest SOA (42 ms) than at the longest SOA (0.4 ms), $F(5, 20) = 24.0, p < .001$ (see Figure 8).

Task 2 absorption prediction. As predicted by bottleneck models, the effect of Task 2 contrast on RT2 was greater at the longest SOA than at the shortest SOAs, $F(5, 20) = 9.0, p < .001$; as shown in Figure 8, the contrast effect was 23 ms at SOA 800, but only −0.2 ms at SOA 100, 2 ms at SOA 17, and 13 ms at SOA −33.

Locus of the bottleneck. If response selection is carried out at or after the bottleneck, then the Task 2 S-R mapping variable should have roughly the same effect on RT2 at all SOAs (i.e., the effects of SOA and compatibility should be additive). The interaction between Task 2 S-R compatibility and SOA was in fact roughly additive, $F(5, 20) = 0.94, p > .20$; the compatibility effect was 12 ms at the longest SOA and 13 ms at the shortest SOA (see Figure 8). This result, taken at face value, suggests that Task 2 response selection was carried out after the bottleneck. Note, however, that the compatibility effect has become very small, greatly reducing our ability to detect any underadditivity. In the General Discussion section, we speculate as to why Experiment 3 might have produced additivity even though Experiment 1 produced underadditivity (i.e., absorption of the factor effects into cognitive slack).

Task 2 error rates. Participants made an average of 5.1% errors on Task 2. This rate decreased with practice, $F(11, 44) = 2.88, p < .01$. There was also a significant interaction between practice and Task 2 S-R mapping, $F(11, 44) = 2.26, p < .05$; participants made more errors with the incompatible S-R mapping early in practice, but this effect went away after a few sessions. In addition, there was a three-way interaction between Task 1 difficulty, Task 2 contrast, and Task 2 S-R mapping, $F(1, 4) = 10.9, p < .05$. This interaction shows no consistent pattern across sessions and was not obtained in the other experiments, so we will not attempt to interpret it.

Task 1 response time. RT1 decreased with practice, $F(11, 44) = 9.3, p < .001$. Except for the main effect of Task 1 difficulty, noted earlier, there were no other significant RT1 effects.

Task 1 error rates. Participants made an average of 8.5% errors on Task 1. As in Experiment 1, Task 1 error rates depended on SOA, $F(5, 20) = 17.2, p < .001$; participants made 12% errors at the shortest SOA and 7.3% errors at the longest SOA. There was also a significant interaction between Task 2 mapping and SOA on Task 1 error rates, $F(5, 20) = 3.14, p < .05$. This reflects a tendency for participants to make a few more Task 1 errors at the shortest SOA when the Task 2 mapping is compatible (13.2%) than when it is incompatible (10.7%).

General Discussion

Van Selst, Ruthruff, and Johnston [VRJ] (1999) showed that extensive practice in a PRP design can dramatically reduce dual-task interference. To learn more about why practice reduces PRP interference, we conducted several follow-up experiments using the VRJ participants. One goal of the present article was to determine why VRJ found a much more dramatic reduction in the PRP effect than had been found by previous investigations. A second goal was to test the bottleneck model with stage shortening (BSS) proposed by VRJ. A further goal was to determine if some of the response selection stage had become automatized with practice (i.e., could be carried out in parallel with central operations on Task 1).

Role of Response Modalities

VRJ speculated that previous practice studies failed to find dramatic PRP reduction because they required manual responses to both tasks, resulting in response conflicts that persisted with practice. To test this hypothesis, Experiment 3 replicated the VRJ experiment with manual responses to both tasks; instead of making a vocal response to Task 1 ("high" vs. "low"), participants pressed one of two responses (the key marked *high* was located just above the key marked *low*). This seemingly minor change caused a very large increase in PRP interference. Whereas the PRP effect found by VRJ was only 50 ms, the initial PRP effect in Experiment 3 was 359 ms. It is interesting to note that this PRP effect is similar in magnitude to that observed in the very first session of the VRJ study (373 ms). Furthermore, whereas the PRP effect declined very rapidly with practice in the VRJ experiment, it showed a modest reduction in Experiment 3. The residual PRP effect after 12 sessions of practice was still about 170 ms, even though by this point the participants had many thousands of practice trials with essentially the same tasks (only the Task 1 response modality changed in Experiment 3). Thus, these results support VRJ's speculation that previous investigators found large PRP effects after practice because they required manual responses to both tasks (Bertelson & Tisseyre, 1969; Borger, 1963; Davis, 1956; Dutta & Walker, 1995; Halliday et al., 1959; Karlin & Kestenbaum, 1968; Van Selst & Jolicoeur, 1997). It is important to note that the elevation in the PRP effects found in Experiment 3, relative to VRJ, cannot simply be attributed to the increase in RT1. As seen in Figure 4, Experiment 3 (diamonds) generally produced a much larger PRP effect for a given RT1 than did VRJ's experiment (plus symbols).

There are several plausible explanations as to why manual-manual designs produce more PRP interference after practice than do vocal-manual designs. One explanation is that the initiation and execution of one manual response interferes with the initiation and execution of any other manual response, even after extensive practice (see De Jong, 1993; Keele, 1973). A related possibility is that response selections on the two tasks suffer heightened cross-talk when their output codes involve the same response modality. Perhaps only when the output modalities are distinct can participants learn to overlap certain central processes, for instance, by selecting responses using two different neural nets, each specialized for a different response modality. Because this hypothesis contends that response selections continue to bottleneck even after practice with a manual-manual design, it can explain why Task 2 S-R compatibility effects were not absorbed into slack at short SOAs in Experiment 3.

A quite different hypothesis is that initiation and/or execution of a manual response to Task 1 simply causes more dual-task interference than the initiation and execution of a vocal response to Task 1 (irrespective of the modality of the Task 2 response). In other words, it might be the Task 1 response modality alone that matters, not the similarity or dissimilarity of the modalities on the two tasks. For example, it is possible that initiation of a manual response requires central resources (prolonging the central bottleneck), whereas initiation of a vocal response does not. According to this hypothesis, the residual PRP effects should be relatively large when the Task 1 response is manual, regardless of the modality of the Task 2 response. If this hypothesis is correct, then large PRP effects should be observed when a manual Task 1 is paired with a vocal Task 2.

Bottleneck Model With Stage-Shortening (BSS)

To account for their results, VRJ proposed the BSS model, which is a straightforward extension of the standard bottleneck model to account for the effects of practice. It says that practice shortens the durations of certain stages on Task 1 and Task 2 but does not eliminate the bottleneck. According to the BSS model, the PRP effect at each level of practice is due to a bottleneck delay whose duration is equal to $1A + 1B - 2A - \text{SOA}_{\text{short}}$ (see Appendix). This formula leads to two key predictions: (a) Task 1 practice should reduce the PRP effect (by reducing 1A, 1B, or both); and (b) Task 2 practice should not reduce the PRP effect (reductions in 2B and 2C should have no effect, whereas reductions in 2A, if any, should only increase the size of the PRP effect).

These two predictions were verified in two transfer experiments using the highly trained participants from VRJ's study. The pairing of the old, highly practiced Task 2 with a new Task 1 produced a relatively large PRP effect (Experiment 1), whereas the pairing of the old Task 1 with a new Task 2 produced only a very modest PRP effect (Experiment 2). Further support for the existence of a processing bottleneck in these experiments comes from the confirmation of the Task 1 carry-over and Task 2 absorption predictions.

The present data also support a special subcase of the BSS model in which practice shortens only the central stages, the BCSS model. This model predicts that declines in both the PRP effect and RT1 with practice should be equal to the decline in the duration of Stage 1B with practice. Hence, declines in the PRP effect and RT1 should be equal to each other. The present data and those of VRJ confirm this prediction with a surprising degree of precision (see Figure 4).

Alternative Models of Practice Effects

Although the BSS model provides a tidy account of the present data, it is nevertheless important to consider whether alternative models can account for them equally well. The two models described in the following discussion are the learned-automaticity model and the task-integration model.

Learned-automaticity model. According to the learned-automaticity model, the PRP effect declines with practice because the component tasks become completely automatized, eliminating the processing bottleneck. Because the automatized tasks require few or no cognitive resources, they should not interfere very much with other tasks, even unpracticed ones. The learned-automaticity model can therefore easily explain the small PRP effect found in Experiment 2 (old Task 1/new Task 2): Task 1 had become automatized and therefore interfered very little with the new Task 2. However, this model also predicts that in Experiment 1 (new Task 1/old Task 2) the highly practiced Task 2 should not have suffered much interference from the new Task 1. Contrary to this prediction, the initial PRP effect in Experiment 1 was relatively large (194 ms).

To rescue this model as a sole explanation of PRP reduction with practice, one would have to assume that Task 1 became automatized with practice in VRJ's study but for some reason Task 2 did not. Although the learned-automaticity model can thus be modified to fit the data, there is not much direct support for it. There was no compelling reason to expect automatization of Task 1 but not of Task 2. It is also important to note that whereas the bottleneck model clearly predicts Task 1 carryover and Task 2 absorption after practice, there is no obvious reason why the learned-automaticity model should make these same predictions. There is also no obvious reason why the learned-automaticity model should predict the almost exact 1:1 relationship observed between declines in the PRP effect and declines in RT1 over the course of practice.

Task-integration model. Even if the component tasks do not become automatized with extended practice, they might nevertheless share the required processing resources in an extremely efficient manner. We refer to this hypothesis as the *task-integration model.* One possibility is that participants learn to treat the two tasks as a single, conjoint task. Alternatively, participants might learn to "tune" or efficiently schedule their mental processes so as to minimize competition for any scarce mental resources. The key assertion of this class of models is that dual-task learning is specific to the particular tasks being practiced together. Thus, whenever either one of the highly practiced tasks is replaced by a novel task, the PRP effect should increase markedly. Performance should then improve as participants learn to integrate the new dual-task ensemble. This prediction matched the outcome of Experiment 1 (new Task 1/old Task 2) but did not match the outcome of Experiment 2 (old Task 1/new Task 2), which produced a small

PRP effect even in the very first session. This result would appear to reject task-integration models.

Task-integration models, however, could be patched by arguing that the new Task 2 judgment in Experiment 2 was similar to the old Task 2 judgment used in the VRJ study. Perhaps participants developed a relatively abstract integration of the two tasks that could be extended to new tasks sufficiently similar to the originals. This patched version of the task-integration model could be disconfirmed by experiments showing strong transfer of learning when there is a greater difference between the old and new Task 2 judgments.

In any case, the task-integration model has other problems as well. This model, like the learned-automaticity model discussed previously, provides no obvious reason to expect Task 1 carryover, Task 2 absorption, or the 1:1 relationship between declines in the PRP effect and declines in RT1 across sessions. Thus, the model cannot yet be definitively ruled out as a sole explanation of PRP reduction, but there is very little direct support for it.

Hybrid Models: Stage Shortening Plus Partial Automatization

We have seen that the BSS model is simple, plausible, and does a good job of explaining both the present data and those from VRJ. The stage-shortening mechanism, by itself, can account for virtually all of the observed PRP reduction. Nevertheless, it is useful to consider whether this was the only change that occurred with practice. It is possible, for instance, that practice also induced partial automatization of the component tasks or that dual-task performance did benefit somewhat from a more efficient integration of the two tasks.

One specific reason to consider such hybrid models is that Task 2 S-R compatibility effects appeared to be partially absorbed into slack at short SOAs in Experiment 1 (new Task 1 / old Task 2). This finding, if taken at face value, indicates that at least some of the Task 2 response selection stage had become automatized with practice (i.e., could operate concurrently with Task 1 central processing). It is not plausible, however, that all of the bottleneck stages were automatized, because even after practice we continued to observe substantial residual PRP effects, Task 1 carryover, and Task 2 absorption. Thus, the data indicate that a bottleneck remained after practice, but that it did not comprise all the same operations that it did prior to practice. Note that this hybrid model is distinct from the learned-automaticity model in that only certain substages have become automatized rather than the entire task.

According to this hybrid model, the main cause of PRP reduction with dual-task practice is Task 1 stage shortening rather than partial automatization. The dramatic decline in RT1 was sufficient to explain virtually all of the decline in the PRP effect with practice in the VRJ study. Furthermore, the stages were so short after practice in the VRJ study that automatization of one or more of the Task 2 substages would likely have contributed relatively little PRP reduction. This assertion is further supported by the data from Experiment 1, in which Task 1 was new and Task 2 was old. That experiment should have benefited from partial automatization of Task 2 (because Task 2 was old), yet it produced a fairly large PRP effect. Furthermore, the modest PRP reduction that was evident in Experiment 1 could be directly attributed to reductions in RT1. Hence, partial automatization of Task 2, by itself, appears to contribute relatively little PRP reduction.

One clue to the nature of partial automatization is provided by our failure to find any evidence of it in Experiment 3, which required manual responses to both tasks. Perhaps automatization is not feasible when response selection on both tasks produces output codes in the same modality, resulting in cross-talk. One specific hypothesis is that automatization is only possible when the response selections on both tasks are subserved by two separate neural nets, each outputting different response codes.

Although we observed evidence for partial automatization with the present tasks, the question of how generally it occurs remains open. The present tasks required a relatively straightforward form of response selection: a one-to-one mapping of a small number of stimuli onto a small number of responses. This arrangement would appear to lend itself well to the establishment of direct S-R associations that are not subject to the central bottleneck. Further research is therefore necessary to determine if the present results extend to tasks that require a more demanding memory retrieval (e.g., where the same stimuli do not occur repeatedly within a session, as in lexical decision tasks) or some difficult and time-consuming mental computation (e.g., multiplication, decision making, or mental rotation).

One specific explanation for partial automatization, which we call the *automatic-activation hypothesis,* says that practice allows participants to select responses based on direct associations between stimulus and response codes. Interestingly, Hommel (1998a) has provided evidence that direct S-R associations, although perhaps rather weak, are established even very early in practice. He found that the speed of the Task 1 response in a PRP design depended significantly on whether that response was compatible with the eventual Task 2 response (see also Logan & Schulkind, 2000). This compatibility effect is intriguing because, according to the central bottleneck model, the Task 2 response cannot be selected until after the Task 1 response has been selected. Consequently, the selection of the Task 2 response should occur much too late to have any influence on RT1.

Hommel (1998a) proposed that although selection and production of the eventual Task 2 response must wait until after selection of the Task 1 response has been completed (due to the central bottleneck), some S-R translation on Task 2 can proceed automatically, causing immediate activation of the correct Task 2 response code. Why cannot participants simply respond based on the automatic activation of response codes and thus bypass the central bottleneck altogether? One possibility is that with low practice levels (the condition studied by Hommel) the automatic response code activation is not strong or reliable enough to permit a sufficiently accurate selection of the Task 2 response. Consequently, participants must instead rely upon a rule-governed or algorithmic response selection that is subject to the central bottleneck (see Logan, 1988, for a discussion of algorithmic vs. automatic response selection). Thus, at low levels of practice, selection of the final response cannot begin until response selection on Task 1 has been completed.

We hypothesize that automatic response activation is relatively weak early in practice, but it becomes much stronger after extensive practice. If so, participants might eventually become able to base their response solely on the automatic activation, without any need for subsequent algorithmic response selection. Note, however, that some further central operation might still be required to take note of whichever response code was most highly activated. Hence, the central bottleneck might not be eliminated entirely, resulting in a residual PRP effect.

Executive Process/Interactive Control (EPIC) Architecture

Meyer and Kieras (1997a, 1997b) have attempted to explain PRP effects using a complex quantitative model in which there is no structural central interference. According to their theory, dual-task slowing instead results from (a) peripheral conflicts in sensory or motor stages or from (b) central postponement that is strategic (i.e., voluntary) in nature. They refer to their model as the *Executive Process/Interactive Control architecture* (EPIC).

Proponents of EPIC might take heart from our finding of partial automatization of the Task 2 response selection stage. They would presumably assert that early in practice (i.e., in VRJ's Session 1) participants voluntarily adopted a cautious strategy of postponing central operations, but after practice participants gradually adopted riskier scheduling strategies that allowed certain central operations to overlap. Thus, they would argue that partial automatization occurred not because practice reduced the resource demands of Task 2 response selection (our hypothesis), but rather because practice induced participants to voluntarily adopt a riskier scheduling strategy.

We have several objections to this account. First, if participants learned a risky scheduling strategy in VRJ's study, then why did a residual bottleneck remain even after practice? Moreover, the full Task 1 carryover indicates that a bottleneck occurred on almost every trial. To rescue EPIC, some new source of interference would presumably need to be invoked.[9] Second, Meyer and Kieras (1997a, 1997b) attributed the cautious scheduling strategy (i.e., voluntary central postponement) primarily to the fact that typical PRP instructions emphasize the speed of Task 1 responses. They suggested that participants might take these instructions to mean "always respond to Task 1 before Task 2." However, in the present experiments and in the VRJ study, these task instructions did not change with practice; hence, participants' task-scheduling strategies should not have changed. Proponents of EPIC might argue that our instructions initially encouraged voluntary postponement, but with practice participants learned on their own that a riskier scheduling strategy was perfectly feasible. If so, these enlightened participants should have continued to use the risky scheduling strategy in the subsequent transfer PRP experiments, resulting in little or no interference. This expectation was not confirmed in Experiment 1, which showed a relatively large PRP effect.

We argue that the willingness of participants to use a risky dual-task scheduling strategy is not the critical ingredient in reducing the PRP effect. This assertion is supported by Ruthruff et al. (in press), who found evidence for central postponement despite instructions to overlap central operations as much as possible (see also Carrier & Pashler, 1995; Ruthruff et al., 1995). So far as we know, the literature contains no instances where unpracticed participants could, by sheer dint of will, perform two demanding speeded response tasks concurrently with little or no interference.

We can summarize the preceding arguments as follows. We found dramatic PRP reduction and partial automatization after extensive practice, despite the use of instructions that Meyer and Kieras (1997a, 1997b) say encourage voluntary postponement. Meanwhile, Ruthruff et al. (in press) failed to find dramatic PRP reduction when instructions strongly discouraged voluntary postponement, but participants were not highly practiced. This contrast suggests that it is task practice, not the instructions or motivation of the participant, that is critical in reducing PRP interference and allowing partial automatization of the response selection stage.

Practical Implications

The present results provide further evidence for VRJ's claim that the PRP effects obtained with novel tasks (usually 300–400 ms) greatly overestimate the amount of interference likely to occur with highly practiced, real-world tasks such as flying, driving, reading, and playing a musical instrument. VRJ found a small residual PRP effect of about 50 ms after 36 sessions of practice. Experiment 1 used a different Task 1 than did VRJ but still produced a 121-ms PRP effect after only four additional sessions of practice; the PRP effect might have dropped even further with additional practice. Likewise, Experiment 2 used a different Task 2 than did VRJ but still produced only a 78-ms PRP effect after eight sessions. It appears, therefore, that the reduction in the PRP effect found by VRJ was not due simply to a fortuitous selection of tasks.

Although we have found several task combinations that produce very small PRP effects after practice, we note that there might be many other combinations that produce much larger residual effects. The bottleneck model clearly predicts that residual PRP effects will be greater for more difficult Task 1 judgments that produce longer RT1s. Furthermore, Experiment 3 makes it clear that residual PRP effects are substantially larger when both tasks require manual responses. Therefore, we caution that firm generalizations about the extent of dual-task interference between highly practiced tasks must await further exploration of the space of possible task combinations.

The present data also have important practical implications for the type of practice necessary to reduce dual-task interference. In Experiment 2, learning transferred very well when the old (highly practiced) Task 1 was paired with a new Task 2. This finding suggests that for training to be effective it is not always critical that the component tasks be practiced together; large training benefits can be obtained even when the tasks are practiced separately. Although training benefits do seem to occur when the tasks are practiced separately prior to being paired together, it is not clear what type of prior practice is most effective. It is possible, for instance, that prior Task 1 practice in a single-task condition would not be nearly as effective as prior Task 1 practice in a dual-task condition (where there is a strong incentive to reduce the resource demands of Task 1).

Summary of Findings

VRJ found that practice dramatically reduced interference in a PRP design. The present research reveals why this finding had not been obtained in previous research: The VRJ result depends on the use of a vocal response to Task 1 and a manual response to Task 2. When both tasks required manual responses (Experiment 3), the

---

[9] Because the VRJ tasks used different output modalities, there should have been no bottleneck in response production (see De Jong, 1993).

PRP effect declined more modestly with practice. Some processing conflict (e.g., in response selection or response initiation) present in the manual-manual design apparently is highly resistant to practice.

Logically, the PRP reduction observed by VRJ could have been due to practice on Task 1, Task 2, or both. The BSS model, however, asserts that Task 1 stage shortening with practice should reduce the PRP effect whereas Task 2 stage shortening should not. Two transfer-of-learning experiments provided strong evidence that practice on Task 1 was indeed much more important than practice on Task 2. Pairing the old, highly practiced Task 2 with a new Task 1 produced a relatively large PRP effect (Experiment 1), whereas pairing the old Task 1 with a new Task 2 produced a small PRP effect (Experiment 2). Furthermore, the nearly exact 1:1 correspondence between declines in the PRP effect and declines in RT1 supports a special case of the BSS model in which practice is assumed to shorten only the central stages.

The absorption of Task 2 S-R compatibility effects into slack (Experiment 1) suggests that practice allowed some of the response selection stage to be carried out prior to the bottleneck. Hence, the data support a bottleneck model in which practice serves both to shorten stage durations and to automatize one or more of the substages that formerly comprised the bottleneck. However, it appears that Task 1 stage shortening, not partial automatization, is responsible for the bulk of the PRP reduction with practice.

Concluding Remarks

Dual-task practice experiments such as those presented in this article are relatively expensive, because training participants takes a long time. Such experiments are also difficult, because practice reduces the size of factor effects and thus weakens the power of statistical tests. Perhaps for these reasons, this area of research has thus far been largely neglected. However, the determination of when and why practice improves dual-task performance is of great practical and theoretical importance. It is also important to determine if performance limitations after practice have the same cause as performance limitations prior to practice. We believe that the work reported here shows that this area of research is tractable and offers rewards commensurate with the efforts required.

References

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.

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology, 105,* 181–199.

Bertelson, P., & Tisseyre, F. (1969). Refractory period of c-reactions. *Journal of Experimental Psychology, 79,* 122–128.

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

Broadbent, D. E. (1982). Task combination and the selective intake of information. *Acta Psychologica, 50,* 253–290.

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

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

Davis, R. (1956). The limits of the "Psychological Refractory Period." *Quarterly Journal of Experimental Psychology, 8,* 24–38.

De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 965–989.

Dutta, A., & Walker, B. N. (1995, November). *Persistence of the PRP effect: Evaluating the response-selection bottleneck.* Paper presented at the 36th annual meeting of the Psychonomic Society, Los Angeles, CA.

Fletcher, B. C., & Rabbit, P. M. A. (1978). The changing pattern of perceptual analytic strategies and response selection with practice in a two-choice reaction time task. *Quarterly Journal of Experimental Psychology, 30,* 417–427.

Gottsdanker, R. (1980). The ubiquitous role of preparation. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in Motor Behavior* (Vol. 3, pp. 301–312). Amsterdam: North-Holland.

Greenwald, A. .(1972). On doing two things at once: Time sharing as a function of ideomotor compatibility. *Journal of Experimental Psychology, 94,* 52–57.

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

Halliday, A. M., Kerr, M., & Elithorn, A. (1959). Grouping of stimuli and apparent exceptions to the psychological refractory period. *Quarterly Journal of Experimental Psychology, 11,* 72–89.

Hirst, W., Spelke, E. S., Reaves, C. C., Caharack, G., & Neisser, U. (1980). Dividing attention without alternation or automaticity. *Journal of Experimental Psychology, General, 109,* 98–117.

Hommel, B. (1998a). Automatic stimulus-response translation in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1368–1384.

Johnston, J. C., & Delgado, D. F. (1993, November). *Bypassing the single-channel bottleneck in dual-task performance.* Paper presented at the 34th annual meeting of the Psychonomic Society, Washington, DC.

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

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

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Logan, G. D., & Schulkind, M. D. (2000). Parallel memory retrieval in dual-task situations: I. Semantic memory. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1072–1090.

McCann, R. S., & Johnston, J. C. (1989, November). *The locus of processing bottlenecks in the overlapping tasks paradigm.* Paper presented at the 30th annual meeting of the Psychonomic Society, Atlanta, GA.

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.

Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of human multiple-task performance: The EPIC information-processing architecture and strategic response deferment model. *Psychological Review, 104,* 1–65.

Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory phenomena. *Psychological Review, 107,* 749–791.

Mowbray, G. H., & Rhoades, M. V. (1959). On the reduction of choice reaction times with practice. *Quarterly Journal of Experimental Psychology, 14,* 1–36.

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

Pashler, H. (1998). *The psychology of attention.* Cambridge, MA: MIT Press.

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

Pashler, H., Carrier, M., & Hoffman, J. (1993). Saccadic eye movements and dual-task interference. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 46*(A), 51–82.

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

Pashler, H., & Johnston, J. C. (1998). Attentional limitations in dual-task performance. In H. Pashler (Ed.), *Attention* (pp. 155–189). Hove, UK: Psychology Press.

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

Ruthruff, E., Pashler, H., & Klaassen, A. (in press). Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement? *Psychonomic Bulletin and Review.*

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

Shaffer, L. H. (1975). Multiple attention in continuous verbal tasks. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 157–167). San Diego: Academic Press.

Smith, M. C. (1969). The effect of varying information on the psychological refractory period. *Acta Psychologica, 30,* 220–231.

Van Selst, M., & Johnston, J. C. (1997). *Dual-task interference when a response is not required.* Proceedings of the 19th annual conference of the Cognitive Science Society, Stanford, CA.

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 47*(A), 631–650.

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

Van Selst, M. A., Ruthruff, E., & Johnston, J. C. (1999). Can practice eliminate the Psychological Refractory Period effect? *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1268–1283.

Welford, A. T. (1952). The "psychological refractory period" and the timing of high-speed performance—A review and a theory. *British Journal of Psychology, 43,* 2–19.

Welford, A. T. (1976). *Skilled performance: Perceptual and motor skills.* Glenview, IL: Scott Foresman.

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

Appendix

The Size of the PRP Effect Predicted by a Generalized Bottleneck Model

The purpose of this appendix is to determine the size of the PRP effect predicted by a generalized bottleneck model (see Figure 1). According to this model, the two tasks can be divided into three serial stages, somewhat arbitrarily labeled A, B, and C. Let $1A$, $1B$, and $1C$ be random variables representing the time required to complete the Task 1 stages, and let $2A$, $2B$, and $2C$ be random variables representing the time required to complete the Task 2 stages.

The key assertion of this model is that Stages A and C can proceed in parallel (without interference) with any other stage, but Stage B (the bottleneck stage) can operate on only one task at a time. Because the Task 1 stimulus generally appears before the Task 2 stimulus, and because participants are told to emphasize the speed of the Task 1 response, we assume that Stage 1B is always processed before Stage 2B (even on occasional trials where Stage 2A finishes before Stage 1A). The onset of Stage 2B might be delayed in some conditions (i.e., at short SOAs); but once it begins, it is assumed to proceed just as quickly as it would if Task 2 were performed in isolation. Note, however that some variable time S (not shown in Figure 1) might be required for participants to switch from performing Stage 1B to performing Stage 2B. Thus, the only sources of dual-task slowing in this model are the bottleneck delay and the switch time.

The PRP effect is generally defined as RT2 at the short SOA minus RT2 at the long SOA:

$$\text{PRP effect} = \text{RT2}_{\text{short SOA}} - \text{RT2}_{\text{long SOA}}. \tag{A1}$$

RT2_{SOA} is equal to $2A + 2B + 2C$ plus any bottleneck delay that occurs at that SOA (see Figure 1). When the bottleneck delay does occur (i.e., when $1A + 1B + S > \text{SOA} + 2A$), the delay will be equal to $(1A + 1B + S) - (\text{SOA} + 2A)$. Otherwise, the bottleneck delay will be zero. Thus, the bottleneck delay is equal to $\max[0, (1A + 1B + S) - (\text{SOA} + 2A)]$. It follows that,

$$\text{RT2}_{\text{SOA}} = 2A + 2B + 2C + \max[0, (1A + 1B + S) - (\text{SOA} + 2A)]. \tag{A2}$$

Substituting this result into Equation A1 gives,

$$\text{PRP effect} = (2A + 2B + 2C + \max[0, (1A + 1B + S) - (\text{SOA}_{\text{short}} + 2A)]) - (2A + 2B + 2C + \max[0, (1A + 1B + S) - (\text{SOA}_{\text{long}} + 2A)]), \text{ and} \tag{A3}$$

$$\text{PRP effect} = \max[0, (1A + 1B + S) - (\text{SOA}_{\text{short}} + 2A)] - \max[0, (1A + 1B + S) - (\text{SOA}_{\text{long}} + 2A)]. \tag{A4}$$

In words, the PRP effect is equal to the bottleneck delay at the short SOA minus the bottleneck delay at the long SOA. At this point, it is worth noting that the durations of Stages 1C, 2B, and 2C appear nowhere in this equation. Consequently, manipulations of these stage durations should have no impact on the size of the PRP effect.

If we assume that the experimenter chooses SOA_{long} so that the bottleneck delay is negligible at that SOA, then this equation reduces to

$$\text{PRP effect} = \max[0, (1A + 1B + S) - (\text{SOA}_{\text{short}} + 2A)]. \tag{A5}$$

If we further assume that $\text{SOA}_{\text{short}}$ is short enough to ensure that there is some bottleneck delay on each trial at that SOA, then Equation A5 further simplifies to

$$\text{PRP effect} = 1A + 1B + S - 2A - \text{SOA}_{\text{short}}. \tag{A6}$$

Note that the latter assumption can be called into question whenever the PRP effect becomes very small. Fortunately, the mean PRP effect was generally greater than about 75 ms in the present experiments, so this assumption is likely to hold reasonably well. Furthermore, if the assumption failed to hold on a small subset of trials, it would complicate the math but would not materially alter any of the key predictions discussed in this article.

The switch time parameter (*S*) has generally played little role in accounts of PRP phenomena. VRJ, for example, found that a bottleneck model with no switch time parameter accurately predicted the 1:1 relationship between declines in the PRP effect and declines in RT1 across sessions. If there were a substantial switch time, then any decrease in this time with practice would have reduced the PRP effect without reducing RT1. Thus, the 1:1 relationship between the PRP effect and RT1 should not have occurred, unless the effects of practice just happened to exactly cancel out some other effect working in the opposite direction. Because there is no apparent need for a switch parameter, and because the bottleneck model is more parsimonious without it, we will not use it in this article. Accordingly,

$$\text{PRP effect} = 1A + 1B - 2A - \text{SOA}_{\text{short}}. \qquad \text{(A7)}$$

Received December 3, 1998
Revision received November 30, 1999
Accepted December 6, 1999 ■

New Editors Appointed, 2002–2007

The Publications and Communications Board of the American Psychological Association announces the appointment of five new editors for 6-year terms beginning in 2002.

As of January 1, 2001, manuscripts should be directed as follows:

- For ***Behavioral Neuroscience,*** submit manuscripts to John F. Disterhoft, PhD, Department of Cell and Molecular Biology, Northwestern University Medical School, 303 E. Chicago Avenue, Chicago, IL 60611-3008.
- For the ***Journal of Experimental Psychology: Applied,*** submit manuscripts to Phillip L. Ackerman, PhD, Georgia Institute of Technology, School of Psychology, MC 0170, 274 5th Street, Atlanta, GA 30332-0170.
- For the ***Journal of Experimental Psychology: General,*** submit manuscripts to D. Stephen Lindsay, PhD, Department of Psychology, University of Victoria, P.O. Box 3050, Victoria, British Columbia, Canada V8W 3P5.
- For ***Neuropsychology,*** submit manuscripts to James T. Becker, PhD, Neuropsychology Research Program, 3501 Forbes Avenue, Suite 830, Pittsburgh, PA 15213.
- For ***Psychological Methods,*** submit manuscripts to Stephen G. West, PhD, Department of Psychology, Arizona State University, Tempe, AZ 85287-1104.

Manuscript submission patterns make the precise date of completion of the 2001 volumes uncertain. Current editors, Michela Gallagher, PhD; Raymond S. Nickerson, PhD; Nora S. Newcombe, PhD; Patricia B. Sutker, PhD; and Mark I. Appelbaum, PhD, respectively, will receive and consider manuscripts through December 31, 2000. Should 2001 volumes be completed before that date, manuscripts will be redirected to the new editors for consideration in 2002 volumes.