Journal of Experimental Psychology:
Human Perception and Performance
1999, Vol. 25, No. 5, 1268–1283



# Can Practice Eliminate the Psychological Refractory Period Effect?

Mark Van Selst
San Jose State University

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

Can people learn to perform two tasks at the same time without interference? To answer this question, the authors trained 6 participants for 36 sessions in a Psychological Refractory Period (PRP) experiment, where Task 1 required a speeded vocal response to an auditory stimulus and Task 2 required a speeded manual response to a visual stimulus. The large PRP effect found initially (353 ms in Session 1) shrank to only about 40 ms over the course of practice, disappearing entirely for 1 of the 6 participants. This reduction in the PRP effect with practice is considerably larger than has been previously reported. The obtained pattern of factor interactions between stimulus onset asynchrony and each of three task difficulty manipulations (Task 1 judgment difficulty, Task 2 stimulus contrast, and Task 2 mapping compatibility) supports a postponement (bottleneck) account of dual-task interference, both before and after practice.

When people are required to respond rapidly to two nearly simultaneous stimuli, substantial interference almost always occurs. Response time (RT) to the second stimulus is often delayed by several hundred milliseconds (Pashler, 1984; Smith, 1969; Welford, 1959). This type of dual-task interference, known as the Psychological Refractory Period (PRP) effect (Craik, 1947, 1948; Hick, 1948; Telford, 1931) or as dual-task slowing (Pashler & Johnston, 1989), has been intensively investigated for many decades because of its practical and theoretical significance. The PRP effect is of practical importance because it demonstrates a severe constraint on the work throughput that can be expected from human operators of complex human–machine systems. The PRP effect is of theoretical importance because it appears to reveal a major limitation in human capability for parallel processing (Allport, 1980; Johnston, McCann, & Remington, 1995; Kerr, 1983; Meyer & Kieras, 1997; Pashler & Johnston, 1998; Welford, 1980), thus providing an important constraint on models of human cognitive architecture (cf. Newell, 1985).[1]

Welford (1959) proposed that the PRP effect is caused by an inability to perform the central operations required for Task 2 (e.g., response selection) at the same time as the central operations required for Task 1. Pashler and Johnston (1989) formalized this model—known as the central bottleneck model—and derived a number of predictions from it. These predictions (discussed below) have since been confirmed in numerous studies (e.g., McCann & Johnston, 1992; Pashler, 1994a, 1994b, 1994c; Pashler & Johnston, 1989; Ruthruff, Pashler, & Klaassen, 1998; Van Selst & Johnston, 1997; Van Selst & Jolicoeur, 1997). Thus, there is compelling evidence that the PRP effect occurs because people do not carry out central operations for two tasks at the same time (see Pashler & Johnston, 1998, for a review).

One of the reasons researchers have devoted so much effort to examining the PRP effect is that it is highly general. It has been shown to occur with many different combinations of stimulus and response modalities (e.g., Fagot & Pashler, 1992; Karlin & Kestenbaum, 1968; McCann & Johnston, 1992; Pashler, 1984, 1994b; Pashler & Johnston, 1989; Ruthruff, Pashler, & Klaassen, 1998; Smith, 1967a, 1967b; Van Selst & Johnston, 1997; Van Selst & Jolicoeur, 1997). It has also been demonstrated using a wide variety of paradigms, including simple, choice, and go/no-go response-time paradigms, and with a wide variety of judgments, including stimulus identification, analog judgments of position and extent, categorization, and naming. Only rarely have exceptions been reported (e.g., Greenwald & Shulman, 1973; Johnston & Delgado, 1993; Halliday, Kerr, & Elithorn, 1959; McLeod & Hume, 1994; Pashler, Carrier, & Hoffman, 1993). All the exceptions appear to involve the special case of tasks in which the stimulus codes might serve

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

This research was supported by fellowships from the National Research Council and NASA Grant NCC2-1043. The three coauthors made equal contributions to the substance of the article.

We thank our participants for their diligence and patience. We also thank Roger Remington, Robert McCann, and Harold Pashler for their insight, thoughts, and commentary.

Correspondence concerning this article should be addressed to Mark Van Selst, 1 Washington Square, Department of Psychology (0120), San Jose State University, San Jose, California 95192. Electronic mail may be sent to mvselst@vision.arc.nasa.gov, eruthruff@mail.arc.nasa.gov, or jcjohnston@mail.arc.nasa.gov.

[1] Another line of dual-task research has used continuous tasks with accuracy as the primary dependent measure (Allport, Antonis, & Reynolds, 1972; Hirst, Spelke, Reaves, Caharack, & Neisser, 1980; Shaffer, 1975). These studies are less useful theoretically because it is difficult to determine whether apparent ability to share tasks successfully does or does not reflect parallel processing. It is virtually impossible to rule out the hypothesis that participants are buffering stimuli and responses and shifting central processing back and forth between the tasks (see McCann & Johnston, 1992, and Pashler & Johnston, 1989, 1998).

as response codes—for example, spatial responses for corresponding spatial stimuli or speech responses that "shadow" spoken words (cf. Greenwald's [1972] ideomotor compatibility hypothesis). In such cases it appears that central operations, such as response selection, can be bypassed (McLeod & Posner, 1984; Norman & Shallice, 1986).

One aspect of the generality of the PRP effect that has not been investigated very extensively, however, is whether it persists with high levels of practice. High levels of practice are commonplace for workers performing a variety of real-world jobs. Many jobs in factories and offices require extensive repetition of a small set of tasks over months or even years. Other jobs, such as aircraft pilot or air traffic controller, are less stereotyped and more cognitively demanding but still involve the execution of highly practiced skills. Before one can apply conclusions from the PRP literature to these important practical domains, it is necessary to first verify that highly practiced tasks are subject to the same dual-task limitations as are relatively unpracticed tasks.

Given the potential importance of practice, it is disappointing and surprising that the vast majority of PRP experiments have—mostly for practical reasons—focused on the performance of participants with low levels of practice. Consider one of the most commonly used laboratory tasks, the tone-frequency discrimination task. Participants classify tones as high or low in pitch and respond by saying "high" or "low" (or by pushing one of two buttons labeled "high" or "low"). Most participants have experience with tone stimuli and with speaking words (or pushing buttons), but they have no practice making these responses to these stimuli. The same is true of most of the other tasks used in PRP studies. Furthermore, it is plausible that what matters most is joint practice on both tasks together (Hirst et al., 1980; Rieck, Ogden, & Anderson, 1980). Because most PRP studies have involved little dual-task practice, one cannot simply assume that the results would be similar with large amounts of dual-task practice.

We have found relatively few studies in the literature for which participants have been practiced over a large number of sessions in the PRP paradigm itself. The few results that have been reported show little evidence that dual-task slowing can be eliminated or even greatly reduced with practice (e.g., Davis, 1956, 1957; Dutta & Walker, 1995; Hick, 1948; Karlin & Kestenbaum, 1968; Van Selst & Jolicoeur, 1997). For instance, in Karlin and Kestenbaum's (1968) often-cited study, a residual PRP effect of 244 ms remained even after extensive dual-task practice. From these results, it would appear that the cognitive limitations responsible for the PRP effect are highly robust across practice.

There are some important reasons, however, to question the generality of this conclusion. Most extended practice PRP studies have chosen task combinations that make it difficult for participants to learn to perform tasks in parallel. In particular, PRP practice studies have almost exclusively used two tasks that both require manual responses (e.g., Bertelson & Tisseyre, 1969; Borger, 1963; Davis, 1956; Dutta & Walker, 1995; Gottsdanker & Stelmach, 1971; Gottsdanker & Way, 1966; Halliday et al., 1959; Karlin & Kestenbaum, 1968; Van Selst & Jolicoeur, 1997). In order to perform two manual tasks in parallel, it would appear to be necessary to control the two hands independently, and motor control research indicates that this is extremely difficult or even impossible (Franz, Eliassen, Ivry, & Gazzaniga, 1996; Franz, Zelaznik, & McCabe, 1991; Tuller & Kelso, 1989).

In fact, PRP experiments at low levels of practice typically show larger interference when both tasks require manual responses than when one of the tasks does not (De Jong, 1993; Fagot & Pashler, 1992; Pashler, 1990; Reynolds, 1964; Van Selst & Johnston, 1997; Vidulich, 1988). One interpretation of this finding is that in addition to the central bottleneck, there is a further bottleneck at the stage of response initiation (De Jong, 1993). If such a response initiation bottleneck were resistant to practice, which seems plausible, then manual–manual tasks would impose a special obstacle to success in reducing the PRP effect with practice. A second possibility is that the use of two similar responses (e.g., in manual–manual designs) increases cross-talk between response selection processes (Wickens, 1991). This cross-talk might inhibit concurrent response selection processes even after extensive practice.

Whatever the reason for the especially large interference observed in manual–manual designs, it seems clear that to observe the effects of practice on central interference one should avoid manual–manual designs. Researchers studying non-PRP dual-task performance (cf. Allport, 1980, and Footnote 1) arrived at a similar conclusion some time ago, but this insight seems not to have been picked up in the PRP domain.

## Goals

The primary goal of this research was to determine the degree to which central interference can be reduced by extended practice. Because our primary interest was in the effect of practice on central processing limitations, we selected tasks that were likely to minimize peripheral stimulus and response conflicts. Each trial required a verbal response to a tone ("high" or "low") and a manual keypress to a visually presented character. This pairing of stimulus and response modalities (auditory–vocal; visual–manual) was found by Shaffer (1975) to minimize interference in a non-PRP dual-task paradigm using accuracy as the dependent measure.

Although our primary goal was to determine the degree to which central interference could be reduced by extended practice, we also wished to determine whether the type of interference changed as a function of practice. More specifically, we were interested in whether the data still showed evidence of a central bottleneck. For this purpose, we factorially manipulated the durations of three different stages in the two tasks. As explained next in some detail, the central bottleneck model makes specific predictions regarding the interactions of these factor effects with stimulus onset asynchrony (SOA). We measured these interactions both before and after extended practice. (For a more

complete justification of these and other predictions of bottleneck models, see McCann & Johnston, 1992; Pashler, 1984, 1992; Pashler & Johnston, 1989, 1998; Schweickert & Boggs, 1984).

## Predictions of Central Bottleneck Models

Consider the central bottleneck model of the PRP effect presented in Figure 1 (after Pashler & Johnston, 1989). The model assumes that processing on each task in the PRP paradigm can be decomposed into three stages, arbitrarily labeled A, B, and C to avoid premature assumptions about the exact nature of each of the three stages.[2] The key assumption of the model is that the two central stages (1B and 2B) cannot operate simultaneously. With that exception, all other combinations of stages from the two tasks are assumed to be able to operate in parallel without interference.

According to the model, processing on Task 1 begins with the presentation of the stimulus, S1, and proceeds through the three processing stages (1A, 1B, 1C), leading to the execution of the response, R1. Processing for Task 2 begins with the presentation of the stimulus, S2, and proceeds immediately through Stage 2A. Stage 2B cannot begin until two preconditions are both met: that Stage 2A has been completed, providing the input for Stage 2B, and that Stage 1B has been completed, making available the processing resources required by Stage 2B. Figure 1 shows the processing diagram for an SOA sufficiently short that Stage 2A is completed before Stage 1B. In this case, a bottleneck has occurred, and the central stage (Stage 2B) is postponed until the Task 1 central stage (Stage 1B) finishes. It is this postponement, or bottleneck delay, that is the cause of the PRP. Once Stage 2B has been performed, the remaining stage (2C) can proceed, leading to execution of the response, R2.

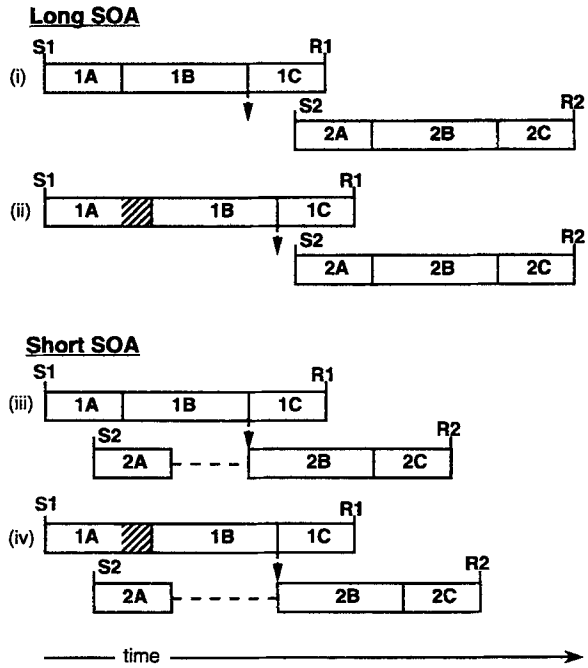


*Figure 1.* Generalized postponement model: Task 1 carryover prediction. Each task involves three stages (A, B, and C). SOA = stimulus onset asynchrony; S1 = Task 1 stimulus onset; S2 = Task 2 stimulus onset; R1 = Task 1 response; R2 = Task 2 response.

Note that the central bottleneck model accounts naturally for the frequent finding of a −1 slope of Task 2 RT (hereafter RT2) across the shortest SOAs, provided that postponement occurs on all trials at these short SOAs (Welford, 1959). Under those conditions, every millisecond of reduction in the SOA adds a millisecond onto the time that Stage 2B must wait for Stage 1B to finish, which in turn adds a millisecond to RT2.

Note that if Stage 1B finishes before Stage 2A finishes, no postponement of Task 2 occurs even though a bottleneck was potentially present (i.e., Stages 1B and 2B could not occur in parallel). We will say that on such no-postponement trials only a latent bottleneck was present. Because of the inevitable variability in the duration of processing stages, any conditions that produce a small mean PRP effect will likely include a substantial subset of latent-bottleneck trials. This proviso turns out to be especially relevant late in practice.

Pashler and Johnston (1989) showed that the central bottleneck model makes several further predictions about the effects of lengthening the duration of various stages of processing of the two tasks (cf. Sternberg, 1969, for a more comprehensive treatment of the theory of processing stages). In the following sections, we briefly review the central bottleneck model predictions that play a key role in the present research.

### *Carryover of Task 1 Difficulty Effects*

The first prediction of the central bottleneck model concerns the effects on RT2 of manipulating the duration of the early or middle stages of Task 1 (see Figure 1; the diagram shows a lengthening of Stage 1A, but the consequences are identical if Stage 1B were lengthened instead). The central bottleneck model makes the strong prediction that whenever bottleneck delays occur, increases in the duration of Stage 1A will not only increase RT1 but will carryover to RT2 as well. This carryover occurs because increases in Stage 1A delay the completion of Stage 1B, which in turn delays Stage 2B, which in turn passes the delay on to RT2. At short SOAs, the model predicts that increases in the duration of Stage 1A will carry over millisecond for millisecond onto RT2, so that RT2 will increase by the same amount as RT1 (this prediction holds exactly only if a bottleneck delay occurs on every trial). At very long SOAs, the model predicts that there will be no carryover of Task 1 manipulations onto RT2 (assuming that in this case postponement never occurs).

[2] Each of these three stages may be decomposed further. The three-stage breakdown, however, is sufficient for the current discussion.

In the experiment reported below, we delayed completion of Task 1 central operations by varying the difficulty of the tone discrimination (note that the carryover prediction is the same whether this Task 1 manipulation affects Stage 1A or 1B). Previous experiments using this manipulation have verified the Task 1 carryover prediction under low levels of practice (Van Selst & Johnston, 1997; see also Smith, 1967a; Williams, 1974).

### *Absorption of Prebottleneck Task 2 Difficulty Effects*

Another prediction of the central bottleneck model (see Figure 2) deals with the effects on RT2 of manipulating the duration of the early, prebottleneck stage of Task 2 (Stage 2A). At long SOAs (where there is no Task 2 postponement), RT2 is simply the sum of the times taken by each of the three component stages. Hence, at long SOAs, an increase of $k$ ms in the duration of Stage 2A should also increase RT2 by $k$ ms. At short SOAs, however, where the bottleneck does occur, the model makes the counterintuitive prediction that an increase of $k$ ms in the duration of Stage 2A will have an effect of less than $k$ ms on RT2. This occurs because during the bottleneck delay, Stage 2B is unable to begin when Stage 2A concludes; hence, a small increase in the duration of Stage 2A will not delay the onset of Stage 2B (it is waiting instead on the conclusion of Stage 1B). Put differently, the bottleneck delay creates slack in the processing of Task 2, which is available to absorb some or all of the time added to Stage 2A (Schweickert & Boggs, 1984). If the slack is greater than $k$ ms on all trials, then an increase of $k$ ms in the duration of Stage 2A will have no effect at all on RT2.

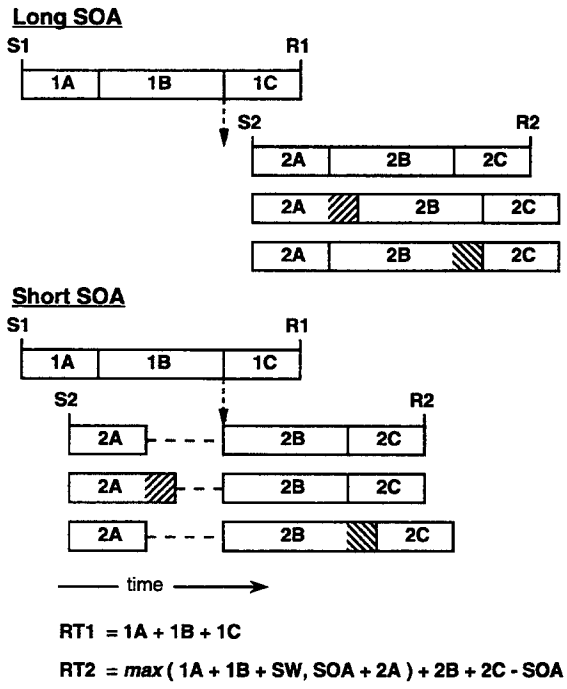


*Figure 2.* Generalized postponement model: Absorption of early-stage Task 2 difficulty effects and additivity of late-stage Task 2 difficulty effects. Each task involves three stages (A, B, and C). The shaded regions represent the additional processing time induced by manipulations of early (Stage A) and central (Stage B) processing demands. SOA = stimulus onset asynchrony; S1 = Task 1 stimulus onset; S2 = Task 2 stimulus onset; R1 = Task 1 response; R2 = Task 2 response; RT1 = Task 1 response time (S1 to R1); RT2 = Task 2 response time (S2 to R2); SW = a parameter representing the time to switch processing from Task 1 to Task 2; max = maximum.

Complete or nearly complete absorption into slack has been confirmed for relatively unpracticed participants when the duration of early visual processing on Task 2 is manipulated by degradation of the stimuli (e.g., De Jong, 1993; Pashler & Johnston, 1989; Van Selst & Johnston, 1997). Confirmation of this counterintuitive prediction has provided some of the strongest evidence for the central bottleneck model. In the experiment below, the duration of early Task 2 processing was lengthened by decreasing the contrast between the visual stimulus (S2) and the background.

### *Additivity of SOA With Manipulations of Task 2 Central Stages*

A further prediction of the central bottleneck model (see Figure 2) deals with the effects on RT2 of manipulating the duration of the central stage of Task 2 (Stage 2B). If there is no postponement (e.g., at long SOAs), RT2 is just the sum of the component stage durations, so adding $k$ ms to Stage 2B will add $k$ ms to RT2. If there is some postponement (e.g., at short SOAs), Stage 2B occurs entirely after the postponement period is over, so there is no opportunity for absorption into slack. Adding $k$ ms to the duration of Stage 2B will again just add $k$ ms to RT2. Thus, the model predicts that the effects of manipulating the duration of Stage 2B will not depend on whether or not postponement occurs and will therefore be the same at all SOAs (i.e., manipulations of SOA and Stage 2B duration should produce additive effects on RT2).

In the experiment we report below, the duration of the Task 2 central stage was manipulated by varying the compatibility of the stimulus–response (S–R) mapping of symbols to response fingers. Previous research on unpracticed participants verified that this type of compatibility manipulation produces effects that are approximately constant across SOA (e.g., Dutta & Walker, 1995; McCann & Johnston, 1992).

### *Declines in the PRP Effect With Practice Should Track Declines in RT1*

Given the strong evidence for the central bottleneck model at low practice levels, it is useful to consider how this model might be extended to account for the effects of practice. Perhaps the simplest extension is to suppose that there are no structural changes in processing; that is, the same component stages continue to exist with practice, and the bottleneck occurs at just the same stages of processing on Task 1 and Task 2 as before. Instead, what practice does is shorten the duration of at least some of the processing

stages. We call this simple model the central bottleneck model with stage shortening.

A plausible subcase of this model asserts that practice primarily reduces the duration of central processing (Fletcher & Rabbitt, 1978; Mowbray & Rhoades, 1959; Pashler & Baylis, 1991; Welford, 1976). Pashler and Baylis (1991) found that single-task practice produces learning that transfers to new S–R mappings if the new mappings use the same classification rule. However, single-task practice fails to transfer to new classification rules, even if the new classification rules use the same stimuli and responses. Pashler and Baylis (1991) concluded from their single-task experiments that practice on ordinary S–R mapping tasks mainly reduces the duration of central processing stages. It is reasonable to assume that practice would have similar effects in dual-task experiments. If so, practice should reduce the duration of the central Stages, 1B and 2B, while having little effect on the noncentral stages.

An interesting prediction follows if one, for the moment, assumes that practice has exactly zero effect on the noncentral stages, 1A, 1C, 2A, and 2C. Suppose that practice reduces the duration of Stage 1B by $k$ ms. This reduction will of course reduce RT1 by the same $k$ ms. Provided that there is a bottleneck delay on every trial, this reduction in the duration of Stage 1B will also reduce the bottleneck delay by $k$ ms and thus will reduce the PRP effect by $k$ ms as well. Because the PRP effect and RT1 are both predicted to decline by the same $k$ ms, the decline in the PRP with practice should track the decline in RT1, millisecond for millisecond. Empirically speaking, a plot of the size of the PRP effect against RT1 across sessions of practice should show a linear relation with a slope of 1.0.

Although practice is thought to primarily influence central processing (Fletcher & Rabbitt, 1978; Mowbray & Rhoades, 1959; Pashler & Baylis, 1991; Welford, 1976), practice may also produce modest decreases in the duration of noncentral stages. Thus, we consider the consequences of the reduction of noncentral stages. Decreases in the duration of Stage 1A will decrease both RT1 and the PRP effect by the same amount and therefore will preserve the anticipated one-to-one PRP:RT1 relationship. Decreases in the duration of Stage 2C reduce RT2 by the same amount at all SOAs, so they would leave the size of the PRP effect unchanged. Because RT1 would also be unchanged, such decreases would have no effect on the PRP:RT1 relationship. Decreases in the duration of Stage 2A would increase the PRP effect, offsetting some of the reduction in the PRP effect that would otherwise occur with practice. Therefore, decreases in Stage 2A with practice would result in a PRP:RT1 slope of somewhat less than 1.0. Decreases in the duration of Stage 1C would also cause the slope to be less than 1.0, because such a decrease would reduce RT1 without altering the amount of PRP interference. In summary, if practice shortens the durations of any noncentral stages, either there will be no effect or else the ratio of PRP reduction to RT1 reduction will fall below 1.0.

It may appear that the prediction of a 1:1 relationship between reductions in the PRP effect with practice and reductions in RT1 with practice depends on so many highly specific assumptions that it stands little chance of actually being true. However, as discussed later, our data showed that the prediction was in fact confirmed with surprising precision.

### *Summary of Predictions*

Central bottleneck models make three distinct predictions for the patterns of factor interactions: (a) Manipulations of Task 1 difficulty will prolong RT1, and this effect will carry over fully onto RT2 at short SOAs but will have little effect at long SOAs; (b) manipulations of prebottleneck Task 2 stages will have a smaller effect at short SOAs than at long SOAs (i.e., will be absorbed into slack); and (c) manipulations of Task 2 central stages will have the same effect at all SOAs.

If the central bottleneck model holds both before and after practice, these predictions should be confirmed both before and after practice. A further prediction—that declines in the PRP effect across sessions should closely track declines in RT1—follows from the more specific central bottleneck model with stage shortening, together with the additional assumption that practice primarily affects the duration of central stages.

## Method

Six participants performed 36 sessions of the experiment (generally one session per day, excluding weekends). The method used in the first 18 sessions (Phase I) is described in detail below. We then describe the minor adjustments made prior to Phase II (Sessions 19–26) and prior to Phase III (Sessions 27–36).

### *Participants*

Each of the six participants had participated in previous PRP experiments. Mark Van Selst and Eric Ruthruff were Participants MV and ER, respectively.[3] The remaining four participants were recruited from work study programs at the National Aeronautics and Space Administration (NASA) Ames Research Center.

### *Stimuli*

*Task 1.* The stimulus for Task 1 was one of four possible tones presented for a duration of 150 ms. The two tones highest in pitch (3125 and 1250 Hz[4]) were classified as high tones, and the two lowest in pitch (200 and 80 Hz) were classified as low tones. Previous research using this Task 1 difficulty manipulation (Van Selst & Johnston, 1997) has shown that in this task, the intermediate high and low tones take longer to classify than the extreme high and low tones.

*Task 2.* The stimulus for Task 2 was an alphanumeric character

[3] The data produced by MV and ER were not qualitatively different from those produced by the other participants.

[4] Low fidelity chassis speakers introduced resonance frequencies and other tonal distortions that might have led to a more difficult tone discrimination than would a discrimination based on pure tones of the same base frequencies.

drawn 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 degrees of visual angle. The background was white; the characters were black (high-contrast condition) or gray (low-contrast condition).

### *Apparatus*

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

### *Procedure*

Participants responded to the pitch of the tone with a vocal response ("high" or "low") and to the identity of the alphanumeric character by pressing the *h, j, k,* or *l* key on a standard keyboard, using the fingers of the right hand. For half of the participants (MV, JC, MR), the letters *A, B, C,* and *D* were mapped in alphabetic order onto the four response keys from left to right (i.e., compatibility), whereas the numbers were mapped in a scrambled order (3, 1, 4, 2) onto the same four response keys, producing an incompatible mapping. For the remaining participants (ER, VL, SW), numbers were mapped compatibly (1, 2, 3, 4) but letters were mapped incompatibly (C, A, D, B). Participants were instructed to respond to both tasks quickly and accurately. Particular emphasis was placed on the speed of Task 1 responses.

Each trial began with the presentation of a fixation cross for 500 ms. The first stimulus followed the offset of the fixation cross by 100 ms. The SOA between the tone, S1, and the alphanumeric character, S2, was 17, 67, 150, 250, 450, or 850 ms.

After each trial, a message appeared if participants made an erroneous response, specifying 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 began with 16 warm-up trials followed by 384 experimental trials. The experimental trials were a random ordering of 2 trials of each of the 192 trial types produced by a complete factorial cross of SOA, Task 1 difficulty, Task 2 contrast, Task 2 compatibility, and Task 2 response finger. Each session was broken into eight blocks of 50 trials each, separated by short breaks. Feedback on the speed of Task 1 and the accuracy of both Task 1 and Task 2 were provided at the end of each block.

### *Methodological Adjustments During Practice*

*Phase II (Sessions 19–26).* The above methodology remained fixed for the first 18 sessions (i.e., Phase I). After Session 18, however, we made three minor methodological adjustments. The first change was to introduce variation into the duration of the interval between the fixation cross and S1 (i.e., the foreperiod). Instead of being fixed at 100 ms, the foreperiod varied randomly from 100 to 250 ms (uniform distribution). This modification reduced the likelihood of Task 1 anticipation errors and anticipation-based processing strategies, which might artifactually attenuate the PRP effect (Gottsdanker, 1979; Karlin, 1959; Van Selst & Jolicoeur, 1997). The second change was to bring the pitches of the two intermediate (i.e., hard) tones closer together: We reduced the 1250 Hz high tone down to 625 Hz and increased the 200 Hz low tone up to 400 Hz. This change was designed to increase the main effect of the tone-task difficulty manipulation, which had become very small by the end of Phase I (making it unlikely that a test of the carryover of this effect onto RT2 would have much power).[5] The final adjustment was to increase S2 contrast in the low contrast condition, which served to reduce the size of the contrast effect. This adjustment was made because, by the end of Phase I, it was clear that the PRP effect after extended practice would otherwise not be large enough to permit full absorption of the degradation effect. A reduction in the size of this effect would, we hoped, allow for a higher proportion of the effect to be absorbed if the central bottleneck model remained valid.

*Phase III (Sessions 27–36).* To further improve our ability to detect and study small amounts of postponement in the final testing phase (Phase III), we reduced all SOAs by 50 ms (from {17, 67, 150, 250, 450, 850} to {−33, 17, 100, 200, 400, 800}). The −33 ms SOA (at which S2 slightly preceded S1 and the 100–250 ms foreperiod was from fixation to S2) was expected to produce a larger PRP effect and therefore provide a cleaner test of the central bottleneck model. The use of a negative SOA might conceivably cause participants to alter their strategy in some way (e.g., they might begin performing central operations on the letter task before central operations on the tone task). However, the negative SOAs were used on only one-sixth of the trials, and the participants already had extensive practice giving emphasis to the tone task. In fact, the data showed no signs of such a strategy shift.

## Results

This experiment produced a rich set of data. To simplify the presentation, we break the results into two sections. In the first section, we describe the decline of PRP interference with practice. In the second section, we present detailed analyses of factor interactions designed to test predictions of the bottleneck model at the beginning of the experiment and after extended practice.

### *The Effect of Practice on the PRP Effect*

Figure 3 shows RT1 and RT2[6] as a function of SOA at four discrete points in time: the first session of Phase I (Session 1) and the final session of Phases I (Session 18), II (Session 26), and III (Session 36). This figure gives the standard snapshot presentation of the data, showing how PRP interference increases as SOA decreases. Figure 4 shows the decline

[5] During Session 1, the tone-task difficulty manipulation increased RT1 by 72 ms and reduced Task 1 percentage correct by 10.2%. By Session 18, the manipulation increased RT1 by only 4 ms and reduced Task 1 percentage correct by only 1.8%. At the start of Phase II (Session 19), the tone-task difficulty manipulation increased RT1 by 28 ms and reduced Task 1 percentage correct by 7.5%.

[6] Response time analyses were restricted to trials in which participants responded correctly to both tasks. Furthermore, a trimming procedure then removed all trials with response times of less than 100 ms, then any trial in which either RT1 or RT2 was identified as an outlier. Outliers were identified using a modified recursive outlier elimination procedure with moving criterion (Van Selst & Jolicoeur, 1994). In Phase I, 3.6% of the data were eliminated from the analyses; 5.9% were eliminated from the Phase II analyses; 5.2% were eliminated from the Phase III analyses.

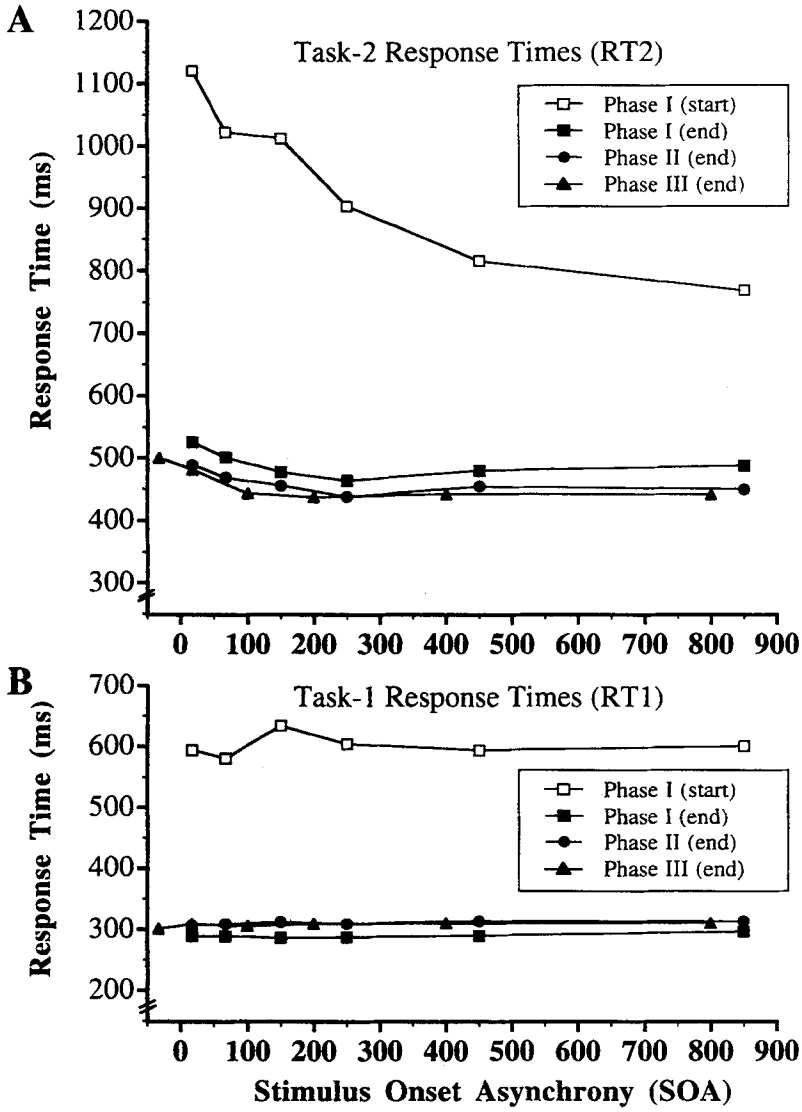


*Figure 3.* The effect of practice on the function relating stimulus onset asynchrony (SOA) and response time for Task 1 (RT1; see Panel B) and Task 2 (RT2; see Panel A).

of PRP interference across sessions of practice when the PRP effect is computed as the difference between the longest SOA and the 17 ms SOA. The solid line in Figure 4 shows the mean PRP effect pooled across the six participants; the dashed lines show the PRP effect for the participants who produced the largest PRP effect (Participant MR) and the smallest PRP effect (Participant SW). The vertical lines demarcate Phases I, II, and III, which differed slightly in method.

*Phase I (Sessions 1–18).* The mean PRP effect at the start of Phase I (i.e., Session 1) was 352 ms (Participant means: MR = 517, MV = 416, ER = 336, VL = 317, JC = 278, SW = 250). As shown in the leftmost panel of Figure 4, the magnitude of the PRP effect decreased rapidly across the first few sessions of practice and then more slowly thereafter.

By the end of Phase I (i.e., Session 18), the average PRP effect was only 38 ms (participant means: MR = 90, MV = 76, ER = 23, VL = 39, JC = 12, SW = −9). This reduction in the PRP effect with practice was accompanied by a substantial reduction in RTs and a modest reduction in error rates.[7] In Phase I, at the longest SOA (the condition in

[7] The overall Task 2 error rate started off at 10.8% in Session 1, dropped to 6.9% in Session 2, then decreased to 3.9% by Session 18. This decrease was significant, $F(17, 85) = 4.36$, $p < .001$. Despite substantial individual differences in overall error rates, each participant showed this downward trend. From Session 1 to Session 18, the error rates for the participants dropped as follows: 22.6% to 5.2% for SW; 13.5% to 5.7% for MV; 11.7% to 3.1% for VL; 7.6% to 5.2% for ER; 4.7% to 1.6% for JC; and 4.4% to 2.6% for MR. Across these same sessions, the overall error rates on Task

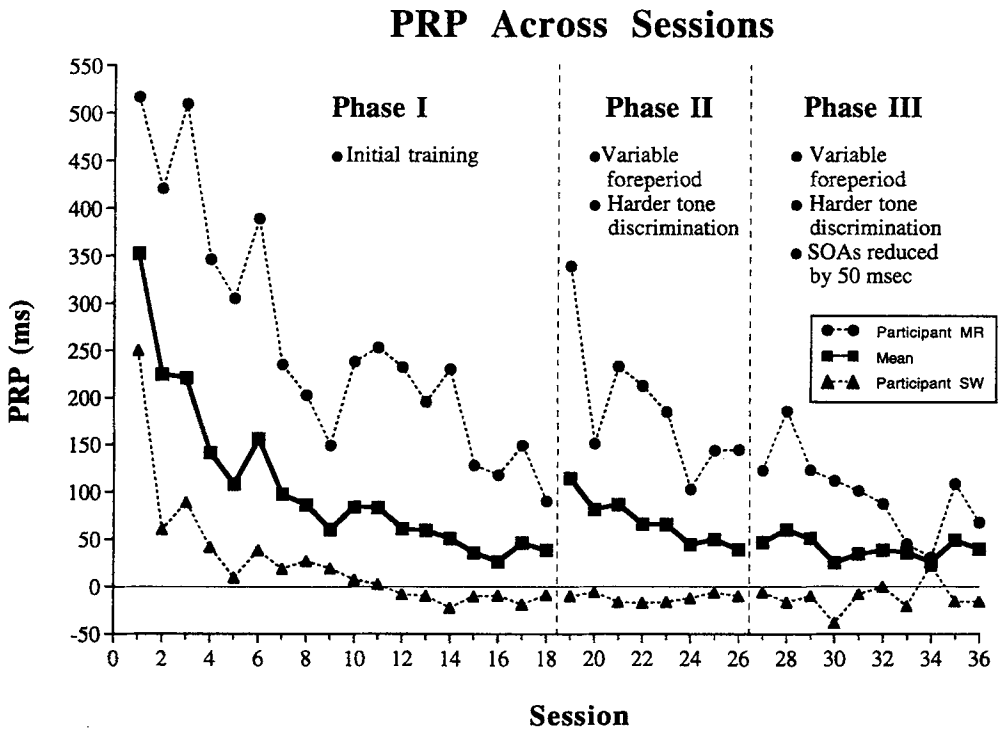


*Figure 4.* Psychological Refractory Period (PRP) effect as a function of sessions of practice. The bold line shows the mean PRP effect pooled across all participants: $RT2_{SOA=17} - RT2_{SOA=850/800}$, where RT2 = response time for Task 2 and SOA = stimulus onset asynchrony. The thin lines show individual PRP functions for the participant who produced the smallest PRP effect (participant SW) and the participant who produced the largest PRP effect (participant MR).

which responses were least likely to be slowed by dual-task interference), RT1 decreased by 304 ms and RT2 decreased by 281 ms.

*Phase II (Sessions 19–26).* At the start of Phase II, we increased the difficulty of the hard tone discrimination required by Task 1. In addition, variability was introduced into the foreperiod between the fixation screen and the onset of S1. These two changes caused the PRP effect to increase at the start of Phase II to 114 ms (Figure 4). After eight additional sessions of Phase II, however, the mean PRP effect had again declined to a low level (39 ms). Broken down by participant, the residual PRP effects in Session 26 were as follows: MR = 145, MV = 25, ER = −2; VL = 42, JC = 36, SW = −10.

*Phase III (Sessions 27–36).* In Phase III, we reduced all of the SOAs by 50 ms. Thus, in Phase III the SOAs were −33, 17, 100, 200, 400, and 800. The shortest SOA in Phase III, the −33 ms SOA, had no counterpart in Phases I and II. Therefore, for the purpose of tracking the PRP effect across phases, we calculated the PRP effect using the shortest SOA common to all phases—the 17 ms SOA. The mean residual PRP effect in Phase III (Sessions 27–36) was only 40 ms, about 11% of its value in Session 1. (Note that this comparison slightly underestimates the effects of practice, because the final Phase III numbers reflect the greater difficulty caused by the random foreperiod and the less discriminable hard tones). Because 17 ms is about as small an SOA as is normally used to measure the PRP effect, 40 ms is the estimate of the PRP effect size that is most appropriate to compare with other published values. The mean Phase III PRP effects (Sessions 27–36), by participant, were as follows: MR = 98, MV = 56, ER = 20, VL = 23, JC = 51, SW = −11.

Using the additional −33 ms SOA condition (used only in Phase III) as a baseline, the mean Phase III PRP effect was 57 ms (individual mean PRP effects were as follows: MR = 130, MV = 83, ER = 22, VL = 51, JC = 68, SW = −9). Note that with this unusually short SOA, it was clear that four of the six participants had residual PRP effects, one was somewhat marginal (ER), and one clearly showed no positive PRP effect (SW).

It is of considerable interest that Participant SW consistently showed no PRP effect late in practice (Figure 4, lower dashed line). If SOA 17 was the shortest SOA for which we had data, we might hypothesize that SW showed no PRP effect because only a latent bottleneck was present on almost all trials. That is, perhaps SW performed Task 1 so rapidly that central Stage 1B now almost always ended before Stage 2A ended. One might expect that this would happen substantially less often at the −33 ms SOA, when Task 2

1 showed a nonmonotonic decrease from 6.3% (Session 1) to 2.5% (Session 18). This decrease failed to produce a significant main effect of session, $F(17, 85) = 1.50, p > .11$.

starts considerably earlier, but the measured PRP effect for SW was less than zero even for that condition. Of course it still remains possible to hypothesize stage durations such that even at the −33 ms SOA there would be no overlap in the central task demands of the two tasks, but much more extreme assumptions would be required. It is noteworthy that late in practice SW consistently showed a small negative PRP effect. The most plausible explanation for this trend was a slightly lower level of preparation for Task 2 at the longer SOAs.

At the beginning of this article, we noted the bottleneck model predicts that the function relating the PRP effect to SOA should have a slope of −1.0, provided that a bottleneck is encountered on every trial. The overall PRP data, even excluding the data from Participant SW, showed an increase in RT2 of well less than the 50 ms difference in SOA between SOA 17 and SOA −33. Thus, the slope of the first segment of the PRP curve was considerably flatter than −1.0 (without SW, the obtained mean slope was only −0.5). The most plausible interpretation of this result was that on a substantial fraction of trials, no postponement occurred (at least in the SOA 17 condition), presumably because only a latent bottleneck was present on those trials.

### *PRP Effect as a Function of RT1 Across Sessions 1–18*

In the beginning of the article, we discussed the possibility that practice reduces central stage durations without eliminating the central bottleneck. According to this model, the PRP effect should be linearly related to RT1 across sessions, with a slope of about 1.0. Figure 5 presents the PRP effect ($RT2_{SOA = 17} - RT2_{SOA = 850}$) as a function of RT1 (at the longest SOA only) for Phase I, where most of the decline in the PRP effect took place. Each data point represents the average of all six participants for one of the first 18 sessions. The linear fit was very good, $r^2 = .96$ (with no systematic deviations), and the slope was 1.02, which did not differ significantly from 1.00.

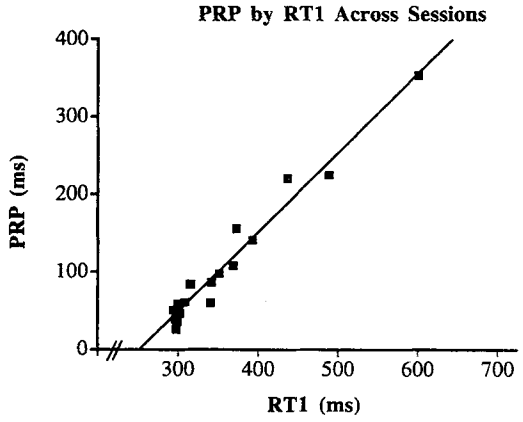


*Figure 5.* Task 1 response times (RT1) as a function of the psychological refractory period (PRP) effect for the first 18 sessions (Phase I). The central bottleneck model predicts a slope of 1.0 for the function relating RT1 and the size of the PRP effect. PRP′ = [(1.022) (RT1)] − 257.342; $r^2 = .959$.

### *Correlation of RT2 and RT1*

The central bottleneck model leads to a very specific prediction about the correlation of RT1 and RT2 across trials. Pashler and Johnston (1989) previously analyzed what the model predicts for relatively unpracticed participants. At short SOAs where the bottleneck occurs, variation in Stages 1A and 1B pushes onto RT2 as well as RT1. Assuming that most of the variation in RT1 is due to Stages 1A and 1B, one should expect substantial positive correlations of RT1 and RT2 (Pashler & Johnston, 1989). On the other hand, at long SOAs no bottleneck occurs, and hence there is no reason for variation in any of stages of Task 1 to push onto RT2. Thus, other things being equal, one would expect RT1 and RT2 to have little correlation.

The prediction of little correlation of RT1 and RT2 at long SOAs should hold late in practice for the same reasons. However, some modification is required in the predictions of the bottleneck model for short SOAs late in practice. Because the mean PRP effect had dwindled to a relatively small size, it seems likely that no actual postponement (only a latent bottleneck) occurred on a substantial fraction of short-SOA trials. For these trials, no positive correlation was expected.

To improve our ability to detect the predicted positive correlations, we restricted our analyses to the data subset most likely to produce postponement. According to the central bottleneck model, the likelihood of postponement is increased by any factor that prolongs Stage 1A or Stage 1B and by any factor that shortens Stage 2A. Hence, we restricted our correlational analysis to the data subset with a difficult Task 1 judgment and an easy (undegraded) Task 2 stimulus. We performed separate analyses for each participant. Higher positive correlations on short-SOA trials should be expected for participants with larger PRP effects and therefore presumably a larger proportion of postponement trials. To ensure a sufficiently large data set, we included trials from the last five sessions of Phase I (Sessions 14–18), where performance was reasonably stable.

At the short SOA, as predicted, RT1 was an excellent predictor of RT2 for those participants with the largest residual PRP effects: $r(32) = .587$, $p < .01$, for MR (*M* PRP effect = 170 ms); $r(42) = .404$, $p < .01$, for MV (*M* PRP effect = 56 ms); and $r(32) = .410$, $p < .02$, for VL (*M* PRP effect = 41 ms). For the participant with the next largest PRP effect, the correlation was marginally significant, $r(36) = .308$, $p < .06$, for ER (*M* PRP effect = 33 ms). For the remaining two participants, who showed very little PRP effect at all, the correlations were small and did not approach significance: $r(32) = -.213$, $p > .20$, for JC (*M* PRP effect = 7 ms); and $r(36) = .046$, *ns*, for SW (*M* PRP effect = 4 ms). At the long SOA, the correlation between RT1 and RT2 was negligible and did not approach significance for any of the participants (*r*s < .11). Thus, even after extensive practice, the data are consistent with the predictions of bottleneck models.

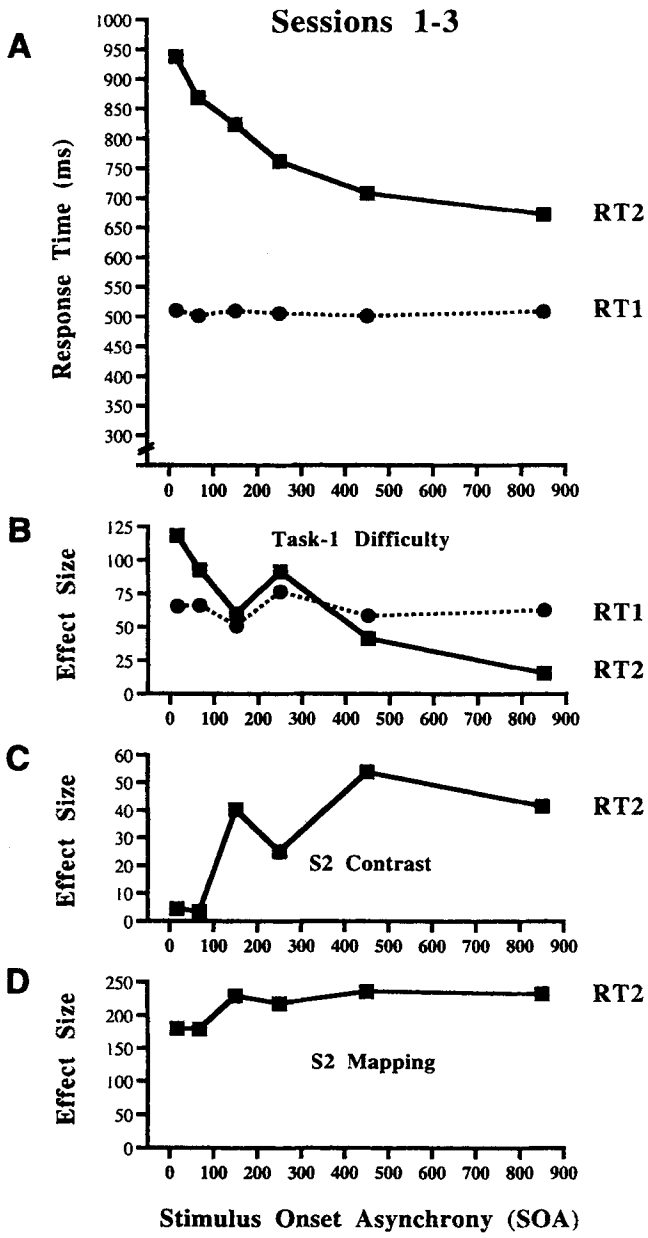


*Figure 6.* Performance measures early in practice (averaged across Sessions 1–3), as a function of stimulus onset asynchrony (SOA). Panel A shows the effect of SOA on Task 1 response times (RT1; represented by dashed line) and Task 2 response times (RT2; represented by solid line). Panel B shows the effect of Task 1 difficulty on RT1 (dashed line) and RT2 (solid line). Panel C shows the effect of Task 2 stimulus onset (S2) contrast on RT2. Panel D shows the effect of Task 2 stimulus–response mapping difficulty on RT2. Note that the scales of the three effect size graphs are not identical.

### *The Interaction Between SOA and Other Experimental Factors*

Performance measures early in practice (Sessions 1–3) are presented in Figure 6 as a function of SOA. These can be compared with the same performance measures after practice (Sessions 27–36), which are presented in Figure 7. For both figures, Panel A shows RT2 (solid line) and RT1 (dashed line). Panel B shows the effect of Task 1 difficulty on RT1 (dashed line) and on RT2 (solid line). Panel C shows

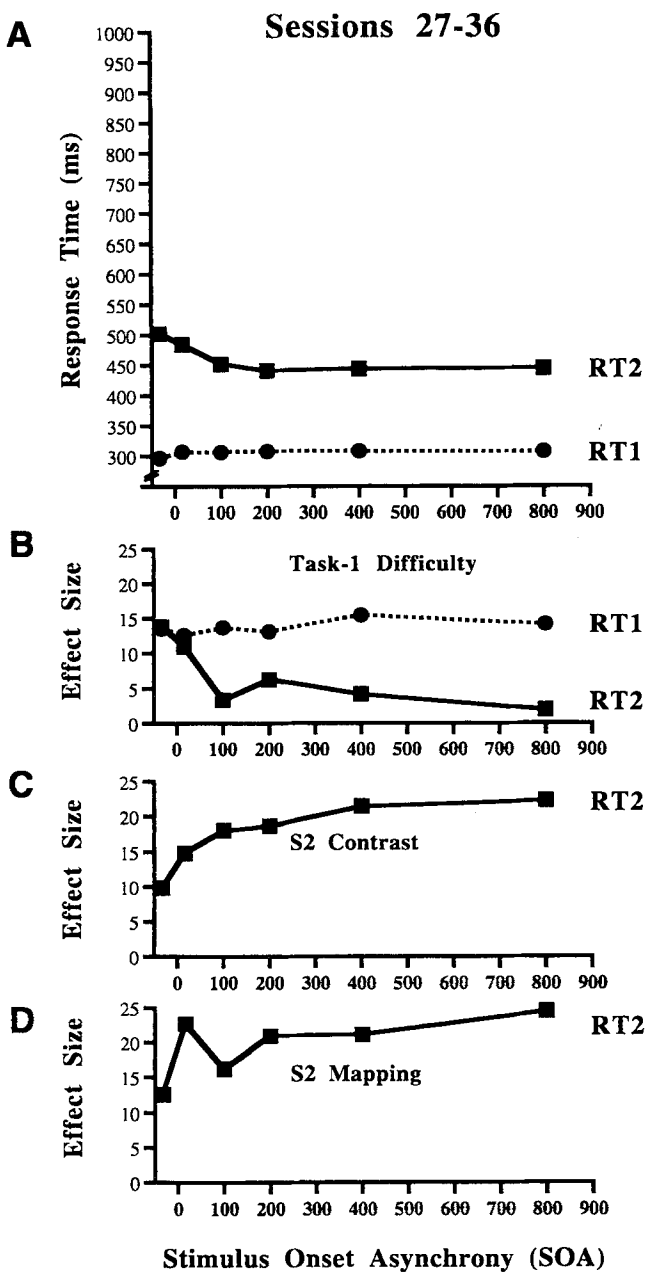


*Figure 7.* Performance measures after extended practice (Sessions 27–36). Panel A shows the effect of stimulus onset asynchrony (SOA) on Task 1 response times (RT1; represented by dashed line) and Task 2 response times (RT2; represented by solid line). Panel B shows the effect of Task 1 difficulty on RT1 (dashed line) and RT2 (solid line). Panel C shows the effect of Task 2 stimulus onset (S2) contrast on RT2 (solid line). Panel D shows the effect of Task 2 stimulus–response mapping difficulty on RT2 (solid line).

the effect of Task 2 contrast on RT2. Panel D shows the effect of Task 2 S–R mapping on RT2.

We performed separate analyses of variance (ANOVAs) on the RT and error rate data for each task, both before and after practice. Omnibus ANOVAs included the full range of six SOAs in addition to the variables of session, Task 1 difficulty (2), S2 mapping (2), and S2 contrast (2). To provide a more sensitive test for whether factor effects differed at long and short SOAs, we also ran similar ANOVAs that were restricted to only the two most extreme levels of SOA.

Performance data from early in practice (see Figure 6) showed a large PRP effect and also confirmed several predictions of the central bottleneck model outlined earlier: (a) The effect of Task 1 difficulty carried over onto RT2 at short SOAs but not long SOAs,[8] (b) the effect of S2 contrast on RT2 was smaller at short SOAs than at long SOAs, and (c) the effect of S–R mapping on RT2 was relatively constant across SOA.[9] In these respects, the data before practice indicated that performance was similar to that of previous PRP experiments with arbitrary S–R mappings.

Late in practice, the pattern of factor interactions was qualitatively similar despite substantial reductions in the size of several of the main factor effects (see Figure 7). For example, the Task 1 difficulty effect on RT1 was 66 ms before practice but only 14 ms after practice. Nevertheless, there was still a clear trend for the Task 1 difficulty effect to carry over at short SOAs (14 ms) but not at long SOAs (2 ms). In the RT2 omnibus ANOVA, the interaction of Task 1 difficulty with SOA was significant, $F(5, 25) = 2.62$, $p < .05$; in the analysis restricted to the longest and shortest SOA, the interaction of Task 1 difficulty and SOA was marginally significant, $F(1, 5) = 6.24$, $p < .06$.

Before practice, reductions in S2 contrast (poorer stimulus quality) slowed RT2 by 42 ms at long SOAs, but after practice this was only a 22 ms effect (almost all of this reduction appeared to be due not to practice, but to the change in method, starting with Phase II, that made the low-contrast condition less difficult). Although the main effect of degradation was smaller, it still decreased with SOA (from 22 ms at the longest SOA to 10 ms at the shortest SOA), as would be expected from absorption into slack. This effect was not quite significant in the omnibus ANOVA using all SOAs, $F(5, 25) = 1.94$, $p < .15$, but was significant in the ANOVA restricted to only the longest and shortest SOAs, $F(1, 5) = 7.05$, $p < .05$.

Practice produced an especially dramatic reduction in the size of the effect of the Task 2 S–R mapping manipulation. For the initial three sessions, the simple main effect of S–R mapping at the longest SOA was very large (232 ms)—not surprising given that the difficult condition specified a very difficult (incompatible) mapping function. After practice, participants were only 25 ms slower with this difficult mapping than with the easy mapping (see Figure 4). This dramatic reduction was consistent with previous evidence that practice serves primarily to decrease the duration of the response-selection stage (Pashler & Baylis, 1991).

The central bottleneck model predicts that the effect of the S–R mapping manipulation should be approximately additive across SOA (cf. McCann & Johnston, 1992). This prediction was only roughly correct, even early in practice. For Sessions 1–3, the S–R mapping effect decreased from 232 ms at the longest SOA to 180 ms at the shortest SOA, $F(5, 25) = 1.63$, *ns*. Although this underadditivity was not significant, it did amount to 23% of the S–R mapping main effect.

Because practice caused such a dramatic reduction in the main effect of S–R mapping difficulty, we were left with a rather noisy test of whether the S–R mapping effect was additive with SOA after practice. The data showed a trend toward underadditivity (12.5 ms at the shortest SOA vs. 25.0 ms at the longest SOA), but the trend was not monotonic. Furthermore, the trend was significant neither in the omnibus analysis, $F(5, 25) = 1.98$, $p < .12$, nor in the analysis restricted to the shortest and longest SOAs, $F(1, 5) = 4.11$, $p < .10$ (see Figure 7). Because some underadditivity was also present in the data early in practice, we were unable to determine whether practice altered the degree to which the S–R mapping stage might show absorption into a later response initiation bottleneck.

## General Discussion

The goal of this research was to determine the extent to which people can learn to perform two independent arbitrary S–R mapping tasks without interference. To answer this question, we trained six participants for 36 sessions each in a Psychological Refractory Period experiment with one vocal response task and one manual response task. Contrary to the results from previous studies (e.g., Van Selst & Jolicoeur, 1997), the PRP effect declined dramatically, dropping from about 350 ms prior to practice to about 40 ms after extended practice. Thus, practice produced nearly a 90% reduction in the size of the PRP effect.

The small size of the PRP effect after practice is especially impressive because we used somewhat more difficult tasks

---

[8] Early in practice (Sessions 1–3), the Task 1 difficulty manipulation slowed RT2 at the shortest SOA even more than it slowed RT1. This result is inconsistent with simple postponement models. However, the pattern was not found later in practice and conflicts with results reported elsewhere (Van Selst & Johnston, 1997); therefore, it might have been artifactual.

[9] Despite previous demonstrations of each of these factor effects and the noise in the data inherent from unpracticed participants contributing only a few observations per cell (eight before removing errors and outliers), some readers may wish to see the statistics for each of these interactions. SOA had a large main effect on RT2, $F(5, 25) = 17.9$, $p < .001$, as did Task 1 difficulty, $F(1, 5) = 13.5$, $p < .001$; S2 contrast, $F(1, 5) = 29.2$, $p < .003$; and Task 2 S–R mapping, $F(1, 5) = 17.2$, $p < .01$. Task 1 difficulty interacted with SOA, $F(5, 25) = 5.96$, $p < .001$. A significant interaction with SOA was found for neither S2 contrast, $F(5, 25) = 1.64$, $p > .19$, nor S–R mapping, $F(5, 25) = 1.64$, $p > .18$. The only shortfall from the predictions described earlier in the article was that the attenuation of the S2 contrast effect, although in the predicted direction, failed to reach significance. Median analyses (arguably more appropriate for noisy data) revealed a similar pattern, with the exception that the interaction between SOA and Task 1 difficulty was now marginal, $F(5, 25) = 2.21$, $p < .09$ ($p < .05$, two-tailed).

than are typically used in PRP studies. In particular, we used more than the usual number of stimuli on both Task 1 (four tones) and Task 2 (eight characters) and more than the usual number of responses on Task 2 (four alternative responses). Our purpose in using more difficult tasks was, in part, to make it unlikely that participants would learn conjoint mapping rules. That is, we wanted to prevent participants from learning to directly associate each possible stimulus pair {S1, S2} with its corresponding response pair {R1, R2} (Kahneman, 1973). In addition to using 32 possible mappings of stimulus pairs to response pairs, we also used a wide range of SOAs, only a few of which were short enough to invite conjoint responding. Furthermore, the SOAs were presented in a random, mixed-list design that made it impossible for participants to know in advance when conjoint responding would be feasible.

Given these precautions, we believe that our highly practiced participants were still performing two separate tasks. Furthermore, because neither task used ideo-motor compatible response mappings (Greenwald, 1972), it seems clear that both tasks required retrieval of response mappings from memory. There was no obvious way for participants to short-circuit the computational requirements of the two tasks, which makes it especially impressive that our tasks could be performed together in the PRP paradigm with so little interference.

### *Relation to Previous Empirical Findings*

Why did previous investigators (e.g., Dutta & Walker, 1995; Karlin & Kestenbaum, 1968; Van Selst & Jolicoeur, 1997) fail to observe a similarly dramatic reduction in the PRP effect across practice? The answer comes in two parts because, compared with previous investigations, we (a) began with a larger PRP effect prior to practice and (b) finished with a much smaller PRP effect after practice.

The large size of our initial PRP effect was at least partly a result of using a more difficult Task 1 than that used in typical PRP practice experiments. Our high–low tone classification Task 1 (with four different tone frequencies) was more difficult than, for example, the "1" versus "2" digit discrimination used by Karlin and Kestenbaum (1968) and by Van Selst and Jolicoeur (1997). The mean RT1 in Session 1 of the current experiment was 601 ms, whereas the mean RT1 in Experiment 2 of Van Selst and Jolicoeur (1997) was only 442 ms. Because our Task 1 was more difficult, it should have resulted in longer postponement of Task 2 central processes and therefore a larger PRP effect prior to practice.

It seems likely that previous investigators found larger residual PRP effects after practice (typically about 250 ms) because participants performed manual responses on both tasks (see Van Selst & Johnston, 1997; Vidulich, 1988). Larger residual PRP effects in manual–manual paradigms might be due to a response-initiation bottleneck (De Jong, 1993). Because practice primarily reduces the duration of central processing (Fletcher & Rabbitt, 1978; Mowbray & Rhoades, 1959; Pashler & Baylis, 1991; Welford, 1976), it seems likely that central interference would decrease with practice, increasing the relative importance of secondary limitations, such as a response initiation bottleneck (De Jong, 1993; Keele, 1973) or peripheral motor constraints (e.g., Meyer & Kieras, 1997). Our use of vocal responses on one task and manual responses on the other should have minimized interference in response initiation and execution. It should also have served to reduce cross-talk between response selections, making it easier for participants to perform them in parallel (provided that no other capacity limitations were present). For instance, with practice one might be able to perform response selections with two separate neural nets, each specialized for a different response modality code.

Another possibility is that interference is elevated with manual–manual PRP designs not because both tasks require manual responses but simply because a manual response is required for Task 1. Regardless of the response modality of Task 2, a manual Task 1 response might cause more interference than a vocal Task 1 response (see Schvaneveldt, 1969; Van Selst & Johnston, 1997). Why might the modality of Task 1 matter? One possibility is that the control of manual responses actually involves spatial attention to finger position, interfering with the spatial processing of S2, whereas vocal responses do not require spatial attention. Another possibility is that manual tasks tend to require more central processing than vocal tasks.

The design of our experiment does not provide any way to discriminate between these possibilities. We believe, however, that the response modality issue is critically important to a more complete understanding of dual-task processing limitations (see Meyer & Kieras, 1997, for a similar view).

### *Can the PRP Effect Be Entirely Eliminated Through Practice?*

One participant (SW) showed no trace of a PRP effect after Session 12. The correct theoretical interpretation of the lack of a PRP effect for Participant SW is not clear. One possibility is that SW learned to perform central operations in parallel, allowing hypothesized central Stages 1B and 2B to overlap (Schumacher et al., 1996). Another possibility is that the central stage of Task 1 was completed before the central stage of Task 2 was set to begin, so that only a latent bottleneck was present.

The other five participants all showed small but significant PRP effects, even after 36 sessions of practice. Might the PRP effect be eliminated altogether for these five participants with further practice? On the one hand, inspection of Figure 4 suggests that the size of the PRP effect has leveled off at a non-zero asymptote by the end of the experiment. On the other hand, even a slight drop in the PRP effect of, say, 0.5 ms per session would eventually lead to the complete elimination of the PRP effect. Hence, although we suspect that most participants have true non-zero asymptotic PRP effects, proving this point would require an indefinitely large number of additional sessions.

Note that the size of the residual PRP effect should depend

on the nature of the particular Task 1 judgment (especially the duration of central processing). Other Task 1 judgments might very well produce either larger or smaller residual PRP effects (e.g., Ruthruff, Johnston, & Van Selst, in press).

### *Is the Residual PRP Effect Due to a Bottleneck?*

The bottleneck theory of the PRP was strongly supported by the pattern of factor effects early in practice. In particular, we observed strong Task 1 carryover effects and absorption of Task 2 contrast effects. The question we now wish to address is whether the small *residual* PRP effect at the end of Phase III was also due to a processing bottleneck.

The data after practice showed many of the usual empirical signs of a processing bottleneck (for a more detailed discussion of these indicators and why they follow from bottleneck models but not from competing models, see Pashler & Johnston, 1998). Perhaps the most impressive evidence is the carryover of the Task 1 difficulty effects onto RT2 at short SOAs (see Figure 7). This finding is a fairly direct indication that delaying Task 1 delays Task 2. Furthermore, we observed a very nearly one-to-one relationship between the decline in RT1 and the decline in the PRP effect across sessions (Figure 5). As discussed above, this one-to-one relationship follows directly from bottleneck models in which practice primarily serves to shorten the duration of central stages: The earlier Task 1 central operations finish, the less time Task 2 central operations must wait. In addition, we found that the effect of S2 contrast decreased as SOA decreased, which reflects the absorption into slack of early-stage Task 2 factors predicted by central bottleneck theory. In summary, we have several converging lines of evidence supporting the conclusion that the residual PRP effect is due to a processing bottleneck.

The conclusion that a bottleneck was present after extended practice does not necessarily imply that postponement occurred on every trial. In fact, that would be extremely unlikely given the very small duration of the mean PRP effect after practice together with the unavoidable stochastic variability in stage durations. Empirical support for probablistic postponement comes from (a) the less than $-1.0$ slope of RT2 against SOA even for the segment between the two shortest SOAs, (b) the less than complete absorption of S2 difficulty effects, and (c) correlations of RT1 with RT2 that, although statistically significant, are smaller than those usually obtained early in practice. The complete carryover of the Task 1 difficulty effect in our data sample is consistent with the existence of a bottleneck on every trial. However, the obtained carryover is small (14 ms) and sufficiently noisy that it can easily be reconciled with the conclusion that the bottleneck occurs only on some trials.

### *What Is the Locus of the Residual Bottleneck?*

There is considerable evidence that, prior to extended practice, the locus of the bottleneck is in central operations (e.g., response selection and perhaps also stimulus classification). We have just discussed evidence that a brief processing bottleneck also exists even after extended practice. But it cannot be assumed without argument that the precise locus of this bottleneck is in the same processing stages as it was early in practice.

Note that the predictions from bottleneck theory of Task 1 carryover and of absorption into slack of early Task 2 factor effects are equally compatible with a central bottleneck locus and a later locus (e.g., at response initiation). The piece of evidence that had the potential to discriminate between these possibilities is the interaction of SOA with our manipulation of Task 2 S–R compatibility. If the bottleneck occurred *at* (or prior to) the stage of response selection, then these effects should have been additive. If there was only a late bottleneck that occurred *after* the stage of response selection, then there should have been an underadditive interaction: The effect of the S–R compatibility manipulation should have been absorbed into the period of cognitive slack caused by the late bottleneck.

It was unfortunate that after extended practice the simple main effect of S–R compatibility at the longest SOA was only 25 ms (compared with 232 ms early in practice), greatly reducing the power of interaction tests. Numerically, the S–R mapping effect after extended practice was smaller at the shortest SOA (13 ms), suggesting some absorption into slack. However, the trend toward a smaller effect at shorter SOAs was not monotonic and was not significant in either the omnibus analysis using all SOAs or the analysis restricted to only the shortest and longest SOAs. Although the effect failed to reach statistical significance, it might nevertheless be genuine. If so, it could mean that after practice the locus of the bottleneck has shifted to a later locus somewhere beyond the stage of response selection (De Jong, 1993; Keele, 1973). On the other hand, it is possible that response selections still bottleneck, but the S–R mapping manipulation after practice influenced not only the duration of the response selection stage, but also the duration of prebottleneck stages, which are subject to absorption into slack.[10]

### *What Caused the Large Drop in the PRP Effect?*

There are several possible hypotheses that could account for the sharp drop in the size of the PRP effect across

[10] Generically, this is known as a failure of the selective influence assumption (Sternberg, 1969). What is in question is the assumption that the manipulation of the S–R mapping rule influences the duration of only the S–R mapping stage. If, for instance, the S–R mapping manipulation had a small effect on the duration of perceptual processes, this portion of the effect of mapping would be absorbed into cognitive slack, producing partial underadditivity between our S–R manipulation and SOA. Early in practice, when the effect of the S–R mapping manipulation on the S–R mapping stage is huge, a small violation of the selective influence assumption would have little noticeable effect (indeed, the number of milliseconds of underadditivity is larger early in practice than late in practice but is a much smaller proportion of the main effect). But late in practice, when participants perform response selection much more rapidly, the small amount of absorption into slack of the effect on perceptual processes could now constitute a substantial proportion of the main effect of the S–R mapping manipulation.

practice. We discuss two hypotheses in turn. The first and arguably the simplest hypothesis is the central bottleneck model with stage shortening, discussed earlier in the article. According to this model, practice does not alter the structural nature of processing for the two tasks: The same central bottleneck model, which does not allow the central stage of Task 1 and the central stage of Task 2 to operate concurrently, accounts for performance both before practice and after practice. Instead, what practice does is drastically shorten the stage durations. Given the findings of Pashler and Baylis (1991), it is reasonable to assume that practice primarily reduces the duration of central stages. It follows from these assumptions that the reduction in the time it takes to complete the central stage of Task 1 with practice would produce a corresponding reduction in the time that Task 2 central processes must wait, which in turn would reduce the PRP effect by the same amount.

A second hypothesis is that participants learn to perform more autonomously some of the central operations of Task 1, Task 2, or both that formerly comprised the processing bottleneck. Thus, a central processing bottleneck might still exist but with fewer processing stages constituting the processing bottleneck.[11] We refer to this possibility as the automaticity hypothesis.

The central bottleneck model with stage shortening provides an appealing and parsimonious account of the great bulk of our data. First of all, the hypothesis has the advantage of invoking a generic assumption—that component stage durations of the two tasks shorten with practice—that virtually must be true because overall RT1 and RT2 drop dramatically with practice. The stage-shortening hypothesis also gains considerable support from successfully predicting that declines in the PRP effect should closely track declines in RT1. Indeed, the relationship between mean RT1 and the size of the PRP effect was found to be almost exactly one-to-one over the first 18 sessions (see Figure 5). The central bottleneck model with stage shortening, because it assumes that a bottleneck exists before and after practice, also provides a well-motivated account for the results of our various factor-effect tests.

What about the automaticity hypothesis? The only positive evidence for this model is the partial underadditivity of the S–R mapping manipulation and the SOA variable after practice. This finding is consistent with some ability to perform the central operations of Task 2 during the bottleneck, although other accounts are also possible. It is difficult to believe, however, that this model accounts for the entire drop in the PRP effect. First of all, it would be curious if the drop in the PRP effect had nothing to do with the large observed drop in RTs for the two tasks. Furthermore, we can find no principled reasons why the automaticity hypothesis should predict the almost exact one-to-one relationship observed between the PRP and RT1. In fact, under the automaticity hypothesis, the entire effect of automization of the central stages of Task 2 would reduce the PRP effect without altering RT1. Thus, if only automatization were at work, we should have observed a much larger drop in the PRP effect than in RT1, rather than the one-to-one relationship actually observed.

What about a hybrid hypothesis that the drop in the PRP is due to both a shortening of central processing stages and automatization? There is no prima facie incompatibility between the two hypotheses. However, the stage-shortening hypothesis would appear to have to account for the bulk of the change in the PRP effect because (a) the tasks *do* shorten with practice, and only the stage-shortening mechanism accounts for this result; and (b) the stage-shortening mechanism contributes toward reducing the PRP millisecond for millisecond with RT1, whereas the automaticity hypothesis does not. There is not much room for a large effect of automaticity without undoing the relationship built by the other mechanism. We should also note that the only positive evidence for the automaticity hypothesis—the not-quite significant underadditivity of the S–R mapping and SOA on RT2 after practice—would be consistent with an automaticity effect of only a few tens of milliseconds, whereas the observed drop in the PRP effect is hundreds of milliseconds.

Overall, it would appear most reasonable to conclude that the great bulk of the reduction in the PRP is produced by stage shortening. There is also a possibility of a modest effect of automaticity, for which the data provide uncertain support. Further experimentation is necessary to pin down whether automaticity does or does not play a role and to test the generality of our conclusions in other practice experiments with different tasks.

### *Practical Importance of the Decrease in the PRP Effect With Practice*

Given that people are well-practiced at many of the tasks they perform in day-to-day life (e.g., typing at a keyboard, driving a vehicle, controlling an aircraft), dual-task studies involving highly practiced tasks have great practical significance. The present study shows that a performance-limiting bottleneck might have less quantitative impact on expert dual-task performance than previously thought—at least with tasks that can be completed relatively quickly. On the other hand, our data show strong evidence that a stubborn residual bottleneck can remain even after extensive practice. Furthermore, our preferred theoretical explanation of the changes in the PRP with practice—the central bottleneck model with stage shortening—suggests that if Task 1 were more complex, larger amounts of dual-task interference might remain, even after extended practice.

A further practical issue concerns the type of practice that is necessary to produce reductions in dual-task interference. One possibility is that dual-task interference will be reduced only if both tasks are practiced together (cf. Hirst et al., 1980). However, most of the hypotheses considered here (especially the central bottleneck model with stage shortening) predict that practice in a single-task condition should decrease dual-task interference about as much as practice in a dual-task condition. Because the present study investigated

[11] More formally, some stage of Task 1 or Task 2 that was part of the bottleneck pool of resources (i.e., that could only be performed on one task at a time) no longer requires bottleneck resources, at least for some subset of the trials in the experiment.

only the effects of practice on both tasks together, further research is needed to determine exactly what conditions are in fact necessary for practice to reduce the PRP effect and other forms of dual-task interference as well.

The present data also touch on several other important practical concerns. We found substantial individual differences in the size of the residual PRP effect after practice. This finding is a tantalizing lead that needs to be followed up to determine what individual traits or capacities determine these differences. Also, a comparison of our results with other results in the literature suggests that extended practice may have little impact on dual-task interference unless different response modalities are used. However, our methods differed from those of previous studies in many ways, so further experimentation is necessary to determine if response modality is in fact the key variable.

### *Summary*

In this study, all six participants showed dramatic reductions in dual-task interference across 36 sessions of practice. One participant showed no sign of a PRP effect after practice, whereas the other five participants showed small residual PRP effects. The pattern of factor interactions and the close correspondence between RT1 and the PRP effect across sessions support the conclusion that a processing bottleneck was present throughout practice. Further work is needed to determine whether the locus of the residual processing bottleneck after practice is central (as appears to be the case at low levels of practice) or is in some later process, such as response initiation.

## References


Allport, D. A. (1980). Attention and performance. In G. Claxton (Ed.), *Cognitive psychology: New directions* (pp. 112–153). London: Routledge and Kegan Paul.

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.

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.

Craik, K. W. (1947). Theory of the human operator in control systems: I. The operator as an engineering system. *British Journal of Psychology, 38,* 56–61.

Craik, K. W. (1948). Theory of the human operator in control systems: II. Man as an element in a control system. *British Journal of Psychology, 38,* 142–148.

Davis, R. (1956). The limits of the "psychological refractory period." *Quarterly Journal of Experimental Psychology, 8,* 24–38.

Davis, R. (1957). The human operator as a single channel information system. *Quarterly Journal of Experimental Psychology, 9,* 119–129.

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.* Poster presented at the 36th annual meeting of the Psychonomic Society, Los Angeles, CA.

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

Fletcher, B. C., & Rabbitt, 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.

Franz, E. A., Eliassen, J. C., Ivry, R. B., & Gazzaniga, M. S. (1996). Dissociation of spatial and temporal coupling in the bimanual movements of callosotomy patients. *Psychological Science, 7,* 306–310.

Franz, E. A., Zelaznik, H. N., & McCabe, G. (1991). Spatial topological constraints in a bimanual task. *Acta Psychologica, 77,* 137–141.

Gottsdanket, R. (1979). A psychological refractory period or an unprepared period? *Journal of Experimental Psychology: Human Perception and Performance, 5,* 208–215.

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

Gottsdanker, R., & Way, T. C. (1966). Varied and constant intersignal intervals in psychological refractoriness. *Journal of Experimental Psychology, 72,* 792–804.

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.

Hick, W. E. (1948). The discontinuous function of the human operator in pursuit tasks. *Quarterly Journal of Experimental Psychology, 1,* 36–51.

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.

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

Johnston, J. C., McCann, R. S., & Remington, R. W. (1995). Chronometric evidence for two types of attention. *Psychological Science, 6,* 365–369.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Karlin, L. (1959). Reaction time as a function of foreperiod duration and variability. *Journal of Experimental Psychology, 58,* 185–191.

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

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

Kerr, B. (1983). Preplanning for aimed movements: Disruptions from a preliminary task. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 596–606.

McCann, R. S., & Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experi-*

*mental Psychology: Human Perception and Performance, 18,* 471–484.

McLeod, P., & Hume, M. (1994). Overlapping mental operations in serial performance with preview: Typing. *Quarterly Journal of Experimental Psychology, 47A,* 193–199.

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

Meyer, D. E., & Kieras, D. E. (1997). 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.

Newell, K. M. (1985). Coordination, control, and skill. In D. Goodman, R. B. Wilberg, & I. M. Franks (Eds.), *Differing perspectives in motor learning, memory, and control* (pp. 295–317). Amsterdam: North-Holland.

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

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. (1990). Do response modality effects support multiprocessor models of divided attention. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 826–842.

Pashler, H. (1992). Dual-task interference and elementary mental mechanisms. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV* (pp. 245–264). Cambridge, MA: MIT Press.

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

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

Pashler, H. (1994c). Overlapping mental operations in serial performance with preview. *Quarterly Journal of Experimental Psychology, 47A,* 161–191.

Pashler, H., & Baylis, G. (1991). Procedural learning: I. 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, 46A,* 51–82.

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

Reynolds, D. (1964). Effects of double stimulation: Temporary inhibition of response. *Psychological Bulletin, 62,* 333–347.

Rieck, A. M., Ogden, G. D., & Anderson, N. S. (1980). An investigation of varying amounts of component-task practice on dual-task performance. *Human Factors, 22,* 373–384.

Ruthruff, E., Johnston, J. C., & Van Selst, M. (in press). Why practice reduces dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance.*

Ruthruff, E., Pashler, H., & Klaassen, A. (1998). *Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement?* Unpublished manuscript.

Schumacher, E. H., Glass, J. M., Lauber, E. J., Gmeindl, L., Woodside, B. J., Kieras, D. E., & Meyer, D. E. (1996, November). *Concurrent response-selection processes in multiple task performance.* Poster presented at the 37th annual meeting of the Psychonomic Society, Chicago, IL.

Schvaneveldt, R. W. (1969). Effects of complexity in simultaneous reaction time tasks. *Journal of Experimental Psychology, 81,* 289–296.

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. (1967a). The psychological refractory period as a function of performance on a first task. *Quarterly Journal of Experimental Psychology, 19,* 350–352.

Smith, M. C. (1967b). Theories of the psychological refractory period. *Psychological Bulletin, 67,* 202–213.

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

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30,* 276–315.

Telford, C. W. (1931). Refractory phase of voluntary and associative responses. *Journal of Experimental Psychology, 14,* 1–35.

Tuller, B., & Kelso, J. A. S. (1989). Environmentally-specified patterns of movement coordination in normal and split-brain subjects. *Experimental Brain Research, 75,* 306–316.

Van Selst, M., & Johnston, J. C. (1997). Dual-task interference when a response is not required. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the nineteenth annual conference of the Cognitive Science Society* (pp. 787–792). Mahwah, NJ: Erlbaum.

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

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

Vidulich, M. A. (1988). Speech responses and dual-task performance: Better time-sharing or asymmetric transfer? *Human Factors, 30,* 517–529.

Welford, A. T. (1959). Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quarterly Journal of Experimental Psychology, 11,* 193–210.

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.

Wickens, C. D. (1991). Processing resources in attention. In D. L. Damos (Ed.), *Multiple-task performance* (pp. 3–34). London: Taylor & Francis.

Williams, L. R. T. (1974). Effects of number of alternatives on the psychological refractoriness of an extended movement. *Journal of Motor Behavior, 6,* 227–234.

Received January 27, 1997
Revision received June 15, 1998
Accepted August 14, 1998 ■