

**Dual-Task Interference with Equal Task Emphasis:
Graded Capacity-Sharing or Central Postponement?**

Eric Ruthruff
NASA - Ames Research Center

Harold E. Pashler
University of California, San Diego

and

Eliot Hazeltine
NASA - Ames Research Center

Abstract

Most studies using the Psychological Refractory Period (PRP) design suggest that dual-task performance is limited by a central bottleneck. Because subjects are usually told to emphasize Task 1, however, the bottleneck might reflect a strategic choice rather than a structural limitation. To address the possibility that central operations can proceed in parallel, albeit with capacity limitations, we conducted two dual-task experiments with equal task emphasis. In both experiments, subjects tended to either group responses together or respond to one task well before the other. In addition, the effects of stimulus-response compatibility manipulation on the character task were roughly constant across stimulus onset asynchronies (SOAs). At short SOAs, character-task compatibility effects also carried over onto response times for the tone task. These results contradict models in which subjects share capacity roughly equally between simultaneous central operations. However, they are consistent with the existence of a structural central bottleneck.

Dual-Task Interference with Equal Task Emphasis: Graded Capacity-Sharing or Central Postponement?

When presented with multiple tasks at the same time, each involving a rapid choice of actions, people tend to exhibit substantial interference. In the laboratory, this interference between tasks is often measured using the Psychological Refractory Period (PRP) design, where two different stimuli requiring speeded responses are separated in time by a variable stimulus onset asynchrony (SOA). Even when the tasks are seemingly very easy (e.g., choice response-time [RT] tasks involving trivial perceptual discriminations), the response to the second stimulus is slowed dramatically at short SOAs. This slowing, known as the PRP effect, is very robust – it has been found in numerous experiments using various combinations of tasks, input modalities, and output modalities.

Several different converging measures point towards the conclusion that the PRP effect is due largely to a central bottleneck; that is, as shown in Figure 1, only one central operation takes place at a time (Welford, 1952; for a review, see Pashler & Johnston, 1998, and Lien & Proctor, 2002; for a contrary view, see Meyer et al., 1995). Central operations include response selection and perhaps also memory retrieval (Carrier & Pashler, 1995), memory consolidation (Jolicoeur & Dell'Acqua, 1998), mental rotation (Ruthruff, Miller, & Lachmann, 1995), and lexical processing (McCann, Remington, & Van Selst, 2000), but generally do not include input/output operations such as stimulus identification and response execution.

Although the evidence for a central bottleneck appears to be strong, it has primarily come from PRP experiments, which usually emphasize Task 1 (i.e., the task corresponding to the stimulus that appears first). For example, subjects might be told to “try to respond to Task 1 as fast as possible.” These instructions are intended to encourage subjects to emit their Task-1 response as soon as it has been selected. Absent these instructions, subjects often withhold their Task-1 response until their Task-2 response is also ready and then emit both responses at the same time.

Address editorial correspondence to:

Eric Ruthruff
NASA - Ames Research Center
MS 262-4
Moffett Field, CA 94035
(650) 604-0343

This response grouping strategy (Borger, 1963) is undesirable from the experimenter's point of view because, among other reasons, it makes it difficult to determine when the critical Task 1 operations actually finished.

The emphasis on the speed of Task-1 raises the question of whether the central bottleneck is obligatory (i.e., due to some structural limitation) or merely voluntary. Subjects might have the latent ability to overlap central operations, but instead choose a bottleneck strategy because it is the easiest way to satisfy the Task-1 emphasis instructions (see McLeod, 1977; Meyer & Kieras, 1997a, 1997b; Meyer et al., 1995; Pashler, 1994a; Ruthruff et al., 1995). There is in fact some reason to believe that people have the latent potential to overlap central operations. Several recent studies have shown that Task-2 processing can influence the RT to Task 1 (RT1), sometimes called a backward compatibility effect. Of particular interest are cases in which the backward compatibility effects on RT1 appear to stem from Task-2 stages at or beyond the central stage (Hommel, 1998; Lien & Proctor, 2000; Logan & Schulkind, 2000; for a review, see Lien & Proctor, 2002). Although the evidence is indirect, such effects are consistent with the hypothesis that Task-2 central stages can be carried out (although not necessarily to completion or with adequate levels of accuracy) in parallel with Task-1 central stages.

In addition, a few recent studies have provided evidence for parallel central processing after considerable practice (i.e., over a thousand trials per task; Hazeltine, Teague, & Ivry, 2002; Schumacher, Seymour, Glass, Kieras, & Meyer, 2001; but see also Levy & Pashler, 2001; Ruthruff, Johnston, & Van Selst, 2001; Van Selst, Ruthruff, & Johnston, 1999; Ruthruff, Johnston, Van Selst, Whitsell, & Remington, in press). Schumacher et al., for instance, found very small dual-task costs (< 10 ms) after subjects had performed five sessions of an auditory-vocal task (saying “1”, “2”, or “3” to low, medium, and high-pitched tones, respectively) and a visual-manual task (making a compatible keypress to the spatial position of a circle). A limitation of these studies, however, is that practice tends to dramatically reduce single-task RT; consequently, little dual-task interference would be expected even if a bottleneck were still present (for a discussion of “latent” bottlenecks, see Ruthruff, Johnston, Van Selst, Whitsell, & Remington, in press). Furthermore, even if the central bottleneck was eliminated in these studies, it might reflect the automatization of one or both tasks with practice. For instance, people might develop “jumper-cable” paths directly between stimulus and response-related areas (Johnston & Delgado, 1993). This explanation seems particularly plausible given the simplicity of the

tasks used in these studies – only three stimuli mapped onto three responses (often with high S-R compatibility).

The purpose of the present paper is to determine if parallel processing of central stages is possible when the instructions encourage it, even without thousands of practice trials. There have been a few previous attempts to address this issue empirically. In particular, Ruthruff, Pashler, and Klaassen (2001) found substantial central interference between tasks even when subjects were asked to group responses together, providing an incentive to carry out the tasks concurrently. Those data argue strongly against the possibility that subjects performed both tasks together with no central interference. However, the existence of central interference does not necessarily imply that there was a central bottleneck. A viable alternative possibility is that central operations proceeded in parallel, but very slowly due to capacity limitations (e.g. McLeod, 1977; Kahneman, 1983). The present paper addresses this possibility, which we call graded capacity-sharing.

Graded Capacity-Sharing

According to the graded capacity-sharing model considered in this paper, all stages on the two tasks (including central stages such as response selection) can proceed in parallel. Due to capacity limitations, however, some of these stages might proceed more slowly in dual-task conditions as compared to single-task conditions. In this paper we focus on a version of this model in which capacity sharing occurs only in central stages (see McLeod, 1977), with the input and output stages generally free from capacity limitations. This model can easily explain the great bulk of PRP experiments simply by assuming that most of the available capacity is allocated initially to Task 1. This assumption is certainly plausible given the typical PRP instructions to emphasize Task 1. Two very recent papers (Navon & Miller, 2002; Tombu & Jolicoeur, in press) have thoroughly evaluated similar capacity-sharing models and arrived at similar conclusions.

Unfortunately, few existing data speak to whether or not graded capacity-sharing is possible. One exception, however, is a study by Pashler (1994a) in which subjects were instructed to place equal emphasis on each of two tasks (a tone task and a letter task). In that study, the SOA between the tone and the letter was -1000, -500, 0, 500, or 1000 ms, so that each task was equally likely to come first. Thus, unlike in the PRP design, there was no implicit or explicit encouragement for subjects to strategically impose a voluntary processing bottleneck.

To determine whether central operations were performed in series (as predicted by bottleneck models) or in parallel (as predicted by capacity-sharing models), Pashler (1994a) examined the patterns of response within a trial. If central operations are performed serially, then the response to Task A should generally be emitted well before the response to Task B, or vice versa. That is, the inter-response-intervals (IRIs) should be relatively large. In addition, there might be a modest proportion of trials where subjects group responses together, producing an IRI close to zero. Thus, there should be three distinct patterns of response. On the other hand, if central operations are performed in parallel with roughly equal capacity allocation, then one might expect a broad, unimodal distribution of IRIs centered roughly near zero. The results showed the response patterns predicted by the central bottleneck model, rather than that predicted by graded capacity-sharing.

The data from Pashler (1994a) indicate that a bottleneck can arise even when tasks are given equal emphasis. However, note that both tasks in this study required manual responses (left hand for the tone task, right hand for the letter task). Therefore, it is possible that the bottleneck was in response initiation or execution (De Jong, 1993; Keele, 1973; but see also Sommer, Leuthold, & Schubert, 2001), rather than in central processing.

The Present Study

The present study was designed to determine whether graded capacity-sharing is possible when the design encourages subjects to overlap performance of the two tasks and when each task uses a

distinct response modality. We chose to study tasks with novel (i.e., not highly-practiced) stimulus-response mappings, which clearly require central processing and should produce relatively long response times (ensuring that a central bottleneck, if present, could be easily detected). Experiment 1 used the basic design of Pashler (1994a), with both positive and negative SOAs and equal task emphasis. However, in the present Experiment 1 we used a tone task with a vocal response and an alphanumeric character-identification task with a manual response (see below for details). The use of different response modalities on the two tasks greatly reduces the likelihood of a response initiation/execution bottleneck. Thus, if we still find evidence for a bottleneck, we can more confidently conclude that the locus of the bottleneck is in central processing.

We used two different approaches to test between the central bottleneck model and graded capacity-sharing models. The first approach was to examine the distribution of IRIs across trials. Experiment 1 tests the bottleneck model predictions outlined by Pashler (1994a). Experiment 2 tests these predictions as well as a novel prediction from the bottleneck model (discussed later) regarding the effects of a slight SOA manipulation (-50 or +50 ms) on the IRIs. In brief, this prediction is that SOA will influence which task's central operations are performed first, but should not affect the shape of the underlying IRI distribution.

The second approach to testing these models (used only in Experiment 1) involved manipulating the duration of central processing on the character task using a stimulus-response (S-R) compatibility manipulation. As discussed below, the central bottleneck model makes very specific predictions regarding (a) how this compatibility manipulation will affect the character task itself and (b) how these effects will carry over onto the tone task. The capacity-sharing model makes much less specific predictions for each of these effects, because the outcome depends critically on how capacity is divided between the two tasks. However, it does make specific predictions for the joint values of these two effects. Based on these joint values, it is possible to estimate the allocation of capacity to the two tasks; that is, whether capacity was divided roughly equally (true capacity-sharing), or whether it was allocated entirely to one task or the other (as predicted by the central bottleneck model). The logic behind these IRI and S-R compatibility predictions is described next.

Central Bottleneck Model Predictions

Inter-Response-Intervals. In this section we describe predictions of the central bottleneck model for the time interval between responses at the 0 ms SOA. To be consistent with our definition of SOA (character onset time minus tone onset time), we define the IRI as the time of the character-task response minus the time of the tone-task response. Thus, positive numbers indicate that the tone-task response was emitted first, and negative values indicate that the character-task response was emitted first.

As noted by Pashler (1994a), the central bottleneck model implies that subjects can perform central operations on the tasks in one of two possible orders on each trial (i.e., character-before-tone or tone-before-character). If subjects select a response to the tone task first and promptly execute that response (i.e., without response grouping), then the tone-task response will be emitted well before the character-task response, producing a large positive IRI. The processing time diagrams for this possibility are shown in the left column of Figure 1 (the top panel corresponds to the compatible condition and the bottom panel corresponds to the incompatible condition). Under the simplifying assumptions that a bottleneck is always encountered at the 0 ms SOA and that the output stages on the two tasks are roughly equal in duration, then the IRI on such trials should simply be equal to the duration of central processing on the character task (plus any task-switching time; see Lien, Schweickert, & Proctor, in press). Based on previous studies (e.g., Pashler, 1994b), we can roughly estimate that this stage will take 300-400 ms, on average; this stage time should vary considerably from trial to trial, as does the overall RT. Note that the duration of this central stage will be much shorter when the character-task S-R mapping is compatible rather than

incompatible, so the IRIs should also be much shorter in the compatible condition.

When subjects select responses in the other order – character before tone – and promptly execute responses (i.e., without response grouping), the character-task response will be emitted well before the tone-task response (see the right panels of Figure 1). Given our IRI definition, these trials should produce negative IRI values. Ignoring the negative sign (which merely indicates which response came first), the IRI in this case should be roughly equal to the duration of central processing on the tone task (again, 300-400 ms on average). Because the character-task S-R compatibility manipulation should have little or no effect on the duration of tone-task central operations, in this case we would expect no effect of compatibility on the IRIs.

A third possibility (not shown in Figure 1) is that subjects will perform central operations in one of the two possible orders, but then will emit responses in close temporal synchrony (“response grouping,” Borger, 1963); that is, they will withhold the response selected first until they have also ready to make the response selected second. In this case, the IRI should be close to zero (perhaps within 150 ms). There should still be some variability in IRIs, but this variability should be far smaller than that observed when responses are not grouped (Pashler & Johnston, 1989). In addition, response grouping causes both tasks to be affected equally by compatibility (regardless of the central processing order); therefore, the resulting IRIs should not be sensitive to character-task S-R compatibility.

Putting these predictions together, the distribution of IRIs aggregated across subjects¹ should have three distinct and identifiable components. One component (tone before character without grouping) should have a mode in the 300 to 500 ms range, and should depend strongly on character-task S-R compatibility. Another component (character before tone without grouping) should have a mode in the -300 to -500 ms range and should not depend on character-task S-R compatibility. The third component (corresponding to response grouping) with a mean near zero should have an especially low variance and should not depend on S-R compatibility.

Compatibility Effects on the Character Task. For the character task, letters were mapped compatibly onto responses but digits were not (for half the subjects), or vice versa (for the other half of the subjects). Because S-R compatibility was manipulated within blocks, subjects did not know in advance of each trial whether the character task would be compatible or incompatible. According to the central bottleneck model, the effects of character-task S-R compatibility on the character task itself should be roughly additive with the effects of SOA (see McCann & Johnston, 1992; Pashler, 1984; Pashler & Johnston, 1989; Schweickert, 1978; Schweickert & Townsend, 1989). In brief, the reason is that the start of the central stage is sometimes delayed at short SOAs (due to the bottleneck), but after this delay it proceeds at full speed. Because this stage proceeds at the same speed in all conditions (and is on the critical path), the effect of S-R compatibility stays the same.

Carryover of Compatibility Effects onto Tone-Task RT. The central bottleneck model also makes specific predictions regarding the effects of the character-task S-R compatibility manipulation on the tone task at the short SOA (0 ms). If tone-task central operations are performed before character-task central operations and the responses are not grouped (see the left panels of Figure 1), then tone-task RT should not be influenced by the S-R compatibility manipulation on the character task – i.e., no carryover should occur. A different result is expected, however, when character-task central operations are performed before tone-task central operations (see the right panels of Figure 1). In this case, prolongation of central operations in the character task (due to the incompatible S-R mapping) should delay the start of central operations on the tone task and thus increase tone-task RT as well. In fact, it should increase by roughly the same amount as the character task RT increases – i.e., there should be full carryover of the compatibility effect onto tone task RT. Carryover should also occur when responses are grouped together, even if tone task central operations were performed first (the tone task response

must wait for character task to finish, and will tend to wait longer when the character task is incompatible).

Summary of Central Bottleneck Model Predictions. In summary, the central bottleneck model makes several clear predictions. First, the distribution of IRIs should have three components: tone before character, character before tone, or grouped (see above for detailed predictions regarding the means and variances of these components). Second, character-task compatibility effects on the character task itself should not depend on the SOA. Third, at the 0 ms SOA, these character-task compatibility effects should not carry over onto tone-task RT when the tone-task response is emitted well before the character-task response, but should carry over fully when the character-task response is emitted well before the tone-task response.

Capacity-Sharing Model Predictions

To derive useful predictions from this class of models, it is necessary to first specify in more detail what is meant by “capacity-sharing.” As a starting point, we adopt the model proposed McLeod (1977; also see Tombu & Jolicoeur, in press), which is reasonably well-specified, plausible, and can account for much of the existing PRP phenomena. As in the central bottleneck model, processing is decomposed into three stages: perceptual processing, central processing, and response execution. The perceptual and response execution stages can operate in parallel with any other stages with no capacity limitations (provided there are no conflicts for a particular input or output modality). The central stages can also operate in parallel with other stages, as shown in Figure 2a, but require a share of a fixed, limited pool of resources. We express the capacity allocated to each task as a proportion of the total capacity. Capacity can be allocated to each of two tasks (Y, Z) in any arbitrary combination in order to meet the current task demands, provided that the values are positive and sum to 1 (i.e., there is no unused capacity while a central stage is underway). Capacity can be dynamically reallocated as soon as a new task needs central processing or an old task completes central processing². Figure 2b shows a possible allocation of capacity to central operations on Tasks Y and Z over time; this example shows a case in which central operations on Task Y finish first, at which point Task Z instantaneously receives the entire capacity³. McLeod (1977) did not specify exactly how the rate of central processing depends on capacity, but here we make the most straightforward assumption, which is that rate is proportional to capacity (see also Pashler, 1984). This capacity-sharing model is similar in many respects to the central bottleneck model, except for the critical assumption that central operations can proceed in parallel. In fact, this model includes the central bottleneck model as a special case where capacity allocation is always all-or-none (i.e., [0,1] or [1,0], where the two numbers in brackets correspond to the capacity allocated to Tasks Y and Z, respectively). Thus, this capacity-sharing model can explain existing PRP data (which are generally consistent with a central bottleneck) simply by asserting that PRP instructions encourage all-or-none capacity allocation. The key issue is whether, under instructions that encourage task sharing, we will continue to see all-or-none allocation (i.e., a bottleneck) or instead see evidence of actual capacity-sharing (e.g., allocations such as [0.5,0.5]).

Inter-Response-Intervals. The present instructions, unlike the usual PRP instructions, did not emphasize one task over the other. Subjects were encouraged to respond to both tasks quickly and accurately. In addition, each task was equally likely to appear first. Thus, if capacity sharing is possible, then these instructions would appear to encourage a roughly even split of capacity between the two tasks (later we will consider more complicated possibilities). If so, we should expect a unimodal distribution of IRIs. This distribution should be fairly broad due to variation in the completion times of the two tasks. Given similar mean single-task RTs to the character and tone tasks, one might expect the center of this distribution to be located somewhere near zero (though it could be shifted away from zero if one task generally receives more capacity than the other). Response grouping might also occur on some trials. If the decision to group responses is random (in the General Discussion we consider

more exotic possibilities), then grouping should simply increase the height of the IRI distribution near 0 ms. As noted by Pashler (1994a), on trials where grouping occurs it is very difficult to discriminate capacity-sharing models from bottleneck models; thus, wherever possible, our analyses focus on subsets of the data where response grouping did not occur.

Compatibility Effects. At long SOAs (-1000 and +1000 ms), where the first task can generally be completed prior to the arrival of the stimulus for the second task, each task should receive the entire pool of resources. At the 0 ms SOA, however, capacity is likely to be shared. The reduced capacity should slow the central stages and thus magnify compatibility effects relative to the -1000 and +1000 ms SOAs (see Pashler, 1984).

The exact size of the increase in compatibility effects at the 0 ms SOA depends on the amount of time during which tone-task central operations overlap with character task operations. If there is no temporal overlap, then no increase should occur. If there is temporal overlap, which seems likely given that the RTs to the tone and character tasks were similar, then an increase in compatibility effects should occur. In this case, the size of the increase also depends on the allocation of capacity. As shown in the Appendix, the maximum possible increase in the character-task compatibility effect at the short SOA is $C_{\text{Tone}}/C_{\text{Char}}$ times the long SOA compatibility effect, where C_{Tone} and C_{Char} are the proportions of capacity allocated to the tone task and character task, respectively. For example, with equal capacity sharing [0.5,0.5] and maximum temporal overlap, the 0-ms SOA compatibility effect would be twice the size of the long SOA compatibility effect.

Carryover of Compatibility Effects onto Tone-Task RT. The predicted carryover of character-task compatibility effects onto the tone task also depends on the amount of overlap between the central operations of the two tasks and on the capacity allocation. If there is no overlap, then no carryover should occur. But if there is complete overlap, then the carryover should be complete; that is, the carryover should be equal to the baseline (i.e., long SOA) compatibility effect on character-task RT (see the Appendix for details).

Because temporal overlap and capacity allocation are unknown, the capacity-sharing model considered here does not make specific predictions about the effects of compatibility on the character itself and the carryover onto the tone task, when considered individually. However, it does make specific, testable predictions about these effects when considered jointly. In fact, as discussed in the Appendix, these effects should be linearly related. The reason is that both effects depend on the exact same set of variables (temporal overlap and capacity allocation). Figure 3 shows the predicted carryover effect plotted against the predicted increase in the effect of compatibility on the character task at the short SOA, expressed as a proportion of the baseline (i.e., long SOA) compatibility effect. Predictions are shown for several different capacity allocations between the two tasks. The predictions for each allocation are represented by a line. The exact location of the prediction along this line depends on the degree of temporal overlap: points on the lower-left (indicating no increase in compatibility effects and no carryover) correspond to the case where there is no temporal overlap, and points on the right correspond to cases where there is substantial temporal overlap. Also shown in Figure 3, on the far left, are two points corresponding to cases in which the central operations are processed strictly serially ([0,1] and [1,0]), as predicted by bottleneck models. The dashed line connecting these two points represents possible mixtures (across trials) of the two strictly serial processing orders. Note that each possible capacity allocation is consistent with only a limited range of combinations of compatibility and carryover effects. Thus, by plotting observed values of these two measures on Figure 3 (as we have done for Experiment 1 using the asterisk symbol), we can estimate the capacity allocations used in that experiment⁴.

Summary of Capacity-Sharing Predictions. The capacity-sharing model discussed above predicts a broad, unimodal distribution of IRIs. In addition, the effects of character-task S-R compatibility on the character task itself should be larger at the 0 ms SOA than at the

long SOAs. Furthermore, these compatibility effects should, under some circumstances, carry over onto tone-task RT. Jointly, the effects of character-task S-R compatibility on the character task itself and on the tone task (i.e., carryover) can be used to estimate how capacity was allocated to the two tasks (see Figure 3).

Experiment 1

The goal of Experiment 1 was to test the predictions of the central bottleneck and the capacity-sharing model described above. To ensure that subjects were familiar with the tasks before we began collecting data, the first four blocks (a total of 152 trials) were considered practice blocks.

Method

Subjects. Thirty-two undergraduates at the University of California, San Diego participated in exchange for partial course credit.

Stimuli. Tone stimuli were presented at 100, 800, or 3000 Hz and lasted for 150 ms. Character stimuli (A, B, C, 1, 2, 3), subtending 1.4° horizontally by 1.9° vertically, were white against a black background.

Procedure. Subjects responded to the 100, 800, and 3000 Hz tones by saying “low,” “medium,” and “high,” respectively. The experimenter taped the session and later verified with a spot check that tone-task accuracy was greater than about 90% for each subject. Subjects responded to character stimuli (A, B, C, 1, 2, 3) by pressing either the ‘j’, ‘k’, or ‘l’ keys. For half the subjects, the letters A, B, and C were mapped in alphabetic order onto the three response keys from left to right (i.e., compatibly), but the numbers were mapped in a scrambled order (3, 1, 2) onto the same three response keys, producing an incompatible mapping. For the remaining subjects, numbers were mapped compatibly (1, 2, 3) but letters were mapped incompatibly (C, A, B).

Subjects completed 12 blocks of 38 trials each (2 warm-up trials plus 36 test trials). Three different SOAs were used: -1000, 0, and 1000 ms. Positive SOAs indicate that the tone sounded before the character appeared, whereas negative SOAs indicate that the character appeared before the tone sounded. Subjects were asked to respond quickly and accurately to both tasks. Both tasks were given equal emphasis.

Each trial began with the presentation of a fixation cross for 1000 ms, followed by a blank field for 500 ms, followed by the tone and/or character stimuli. If an error was detected, an error message was displayed for 1 sec. The next trial began 1 sec later. At the end of each block of trials, subjects received performance feedback (RT and percent correct) and were allowed to take a short break.

Results

Subjects with character-task error rates greater than 15% in either the compatible or the incompatible condition were eliminated from the analyses (N=3), as were subjects who tended to group responses even at the -1000 and 1000 ms SOAs (N=3). In addition, trials with either an error, an RT less than 200 ms (< 1% of all trials), or an RT greater than 3 sec (< 1% of all trials) were excluded from the analyses.

Main Effects. Mean tone-task and character-task RTs are shown in Table 1 as a function of SOA and character-task response compatibility; character-task error rates are shown in parentheses. There was a significant main effect of SOA on both the character task, $F(2, 50) = 84.1, p < .001$, and the tone task, $F(2, 50) = 116.9, p < .001$, reflecting substantial dual-task slowing at the 0 ms SOA. There was also a significant main effect of compatibility on the character task, $F(1, 25) = 63.2, p < .001$; RTs were 174 ms longer, on average, in the incompatible condition than in the compatible condition. Error data on the character task showed a small but significant main effect of SOA, $F(2, 50) = 3.7, p < .05$. In addition, errors were much less frequent in the compatible condition (0.4%) than in the incompatible condition (4.9%), $F(1, 25) = 72.4, p < .001$.

IRI Distributions. Figure 4A shows the distribution of IRIs for the 0 ms SOA condition. Plotted are the proportion of responses, averaged across subjects, in each 60 ms bin; the open circles represent the compatible condition and the filled squares represent the incompatible condition. Qualitatively, the aggregate IRI distributions appear to

match the predictions of the central bottleneck model. That model predicted that there would be one high-variance component with a mode between -300 to -500 ms (where tone-task central operations were performed first with no response grouping) that should depend on character-task compatibility and another high-variance component with a mode between 300 to 500 ms (where character-task central operations were performed first with no response grouping) that should not depend on character-task compatibility. There is also apparent evidence for the third predicted component (corresponding to grouped responses) with relatively low variance and a mean near zero.

Note that although three components are evident in the aggregate data, individual subjects often show evidence for only one or two of the predicted components (e.g., some do not group responses, others prefer one response order over the other). This point is evident in Figure 4B, which shows the histogram of IRIs (averaged across the compatible and incompatible conditions) for two sub-groups of population. The majority group (N=18; diamond symbols connected by a solid line) consisted of the subjects who produced the fewest responses near zero, and the minority group (N=8; plus symbols connected by a dotted line) consisted of the remaining subjects, who produced the most responses near zero. The majority group clearly shows the bimodality predicted by bottleneck models; the two modes fall just where the bottleneck model predicted they would fall (plus or minus 300-400 ms). The minority group (suspected to be grouping on a substantial proportion of trials) also shows these same two components in the same places, albeit with a smaller overall proportion of trials. However, this minority group also shows evidence for a narrow spike of trials with an IRI close to 0 ms. The narrowness of this spike is consistent with response grouping (i.e., an effort to synchronize responses). Thus, the data suggest with a bottleneck model and the additional proviso that a minority of subjects frequently grouped responses.

The most straightforward prediction from capacity-sharing models with roughly equal sharing, would be a unimodal distribution of IRIs. This prediction was not confirmed. The dip in response proportion in the bins between 90 and 210 ms was not predicted by this model, nor was the high proportion of responses in extreme positive or negative IRI bins (e.g., beyond -450 or +450 ms).

To provide a quantitative test between these models, we compared the number of responses in the 90-210 ms IRI bins (both positive and negative) to the number of responses in the 270-390 ms bins (both positive and negative). Note that the two bin groups contain the exact same range of IRI values (a total of 240 ms each). We reasoned that capacity-sharing models clearly predict many more responses in the bins closer to zero (90-210 ms) than in the extreme bins (270-390 ms). According to the central bottleneck model, however, there might be relatively few responses in the IRI bins closer to zero (90-210 ms), which are too long for most grouped responses yet too short for most non-grouped responses. Meanwhile, there might be many responses in the extreme bins (270-390 ms), which roughly coincide with the expected mode of the non-grouped responses. Hence, the central bottleneck model can predict the opposite – and perhaps more counterintuitive – result: more responses at the extreme bins than at the bins closer to zero. Note, however, that even if there is a central bottleneck, not every subject would necessarily show this pattern; in particular, subjects who primarily group responses should not show this pattern.

Even though the extreme IRI bins are further from zero, they contained a higher proportion of responses, on average, than the bins closer to zero (0.16 vs. 0.12). Sixteen of the 26 subjects showed this pattern. The ones who did not show this pattern were those that appeared to consistently adopt a response grouping strategy; not coincidentally, these subjects were part of the minority group shown in Figure 4B. Although the difference between these bins was only marginally significant, $F(1, 25) = 2.6$, $p < .1$, the observed means are more consistent with the predictions of the central bottleneck model (plus occasional grouping) than with those of the capacity-sharing model. A primary goal of Experiment 2 will be to replicate these IRI

results in a design that produces less response grouping, yielding cleaner data.

Compatibility Effects X SOA. Consistent with the central bottleneck model, the data showed no significant interaction between compatibility and SOA on the character task, $F(2, 50) = 1.47$, $p > .2$; the effect of compatibility was 160 ms at the -1000 ms SOA, 161 ms at the 0 ms SOA, and 201 ms at 1000 ms SOA. The capacity-sharing model, however, incorrectly predicted an increase in compatibility effects at the 0 ms SOA. Neither model provides an obvious explanation for the somewhat larger compatibility effects at the 1000 ms SOA (where the letter task was performed second) compared to the -1000 ms SOA (where the letter task was performed first); however, this difference was not significant, $F(1, 25) = 2.27$, $p > 0.1$, and therefore might simply be due to sampling error.

Carryover of Response Compatibility Effects onto Tone-Task RT. At the short SOA (0 ms), we observed significant effects of character-task compatibility on the tone task. Overall, this carryover effect was 67 ms. The central bottleneck model predicts no carryover when the tone-task response was emitted well before the character-task response, but predicts full carryover (i.e., roughly equal in size to the compatibility effects on the character task itself) onto tone-task RT when the character-task response was emitted well before the tone-task response. Thus, the 67 ms average carryover effect would reflect a mixture of two states, one producing no carryover and one producing full carryover (as noted above, response grouping also produces full carryover). To evaluate this prediction, trials with an IRI more negative than -210 ms were assigned to the tone-before-character category, whereas trials with an IRI greater than +210 were assigned to the character-before-tone category. Because not all subjects produced adequate numbers of responses in each category to permit a meaningful estimate of carryover, each estimate was based only on data from subjects with at least 5 trials in both the compatible and incompatible conditions. Using this approach, the carryover of compatibility effects onto tone-task RT in the tone-before-character category was -21 ms (based on 17 subjects, with an average of 11.6 compatible trials and 13.1 incompatible trials), but carryover in the character-before-tone category was 170 ms (based on 18 subjects, with an average of 13.2 compatible trials and 11.6 incompatible trials). Thus, we observed no carryover in one case and full carryover (roughly equal to the average effect of compatibility on the character task itself at the 0 ms SOA of 161 ms) in the other, just as predicted by the central bottleneck model.

The carryover of character-task S-R compatibility effects onto the tone task, combined with the effects of compatibility on the character-task itself (discussed in the preceding section), argue against roughly equal capacity-sharing. Figure 3 shows the predicted combinations of values for these two effects as a function of the capacity allocation; the asterisk shows the values observed in Experiment 1 along with standard error bars⁵. The data fall between the predicted results from the all-or-none allocations ([0,1] and [1,0]); hence the data are consistent with a mixture of the two possible orders of sequential central processing predicted by bottleneck models (tone-before-character and character-before-tone). Meanwhile, the data are inconsistent with the predictions from roughly equal capacity allocations between central operations on the two tasks.

Near-Zero IRIs: Grouping or Capacity-Sharing?

Although the data suggest that subjects often chose extreme allocations of capacity to the central stages, consistent with the bottleneck model, it is nevertheless conceivable that capacity-sharing occurred for a substantial subset of trials (or for a subset of subjects). In particular, there were a substantial number of trials with IRIs close zero that might have resulted from roughly equal capacity-sharing. According to the bottleneck model, however, these trials resulted from response grouping.

There are two findings that support this response grouping hypothesis. First, these trials show a very tall, narrow peak (i.e., with low variability), just as one would expect if the responses were produced as a couplet (Pashler & Johnston, 1989). Second, if there is

a central bottleneck and responses are grouped, then the character-task compatible effect should carryover fully onto tone-task RT (regardless of the order of the central operations). Indeed, for IRIs between +150 and -150 ms, the average effect of compatibility on the character task was 123 ms and the average carryover onto the tone task was 116 ms. Because the data with IRIs near 0 are consistent with a central bottleneck plus response grouping, they do not constitute evidence for capacity-sharing.

Discussion

The results were generally consistent with the central bottleneck model. First, the distribution of IRIs were consistent the predicted three-component pattern. The central mode (with a peak near 0 ms IRI) is assumed to correspond to trials in which subjects grouped responses. The two extreme modes (with peaks near +350 and -350 ms) correspond to the two possible orders in which central processes can be executed (without response grouping). A second finding supporting the central bottleneck model concerns the effects of the character-task S-R compatibility manipulation. Specifically, the effects of this manipulation on character-task RT were roughly constant across SOAs. In addition, these effects did not carry over onto tone-task RT when the tone-task response came first, but carried over fully when the character-task response came first.

The results, however, did not match the predictions of the capacity-sharing model outlined in the introduction. According to this model, central operations operate in parallel but must share a limited pool of resources. First, this model predicts a unimodal distribution of IRIs, contrary to our observations (see Figure 4). There was a tendency for subjects to respond to one task well before the other (e.g., by 300-400 ms) or nearly simultaneously. Meanwhile, there were relatively few responses in intermediate IRI bins (e.g., the IRI bins centered at 60 and 120 ms). The capacity sharing model also has difficulty accounting for the observed effects of character-task S-R compatibility. As can be seen in Figure 3, the observed combination of compatibility effects on the character-task itself and on the tone task (carryover) is inconsistent with the capacity-sharing model where capacity is allocated roughly evenly between simultaneous central operations. Instead, it suggests the use of extreme capacity allocations, such as those predicted by bottleneck models ([0,1] and [1,0]).

Experiment 2

Experiment 2 had two main purposes. One purpose was replicate Experiment 1 with less frequent response grouping and another was to test a novel prediction regarding the IRI distributions. We used a variant of a dual-task design that has recently been used by several investigators to determine whether dual-task interference can be eliminated after practice (Hazeltine, Ruthruff, & Remington, 2001; Hazeltine et al., 2002; Levy & Pashler, 2001; Schumacher, Seymour, Glass, Kieras, & Meyer, 2001). In this design, there are three conditions, a dual-task condition and two single-task conditions (one for each task). These single- and dual-task conditions are mixed together within a block, so subjects cannot know which combination of tasks will be presented on the upcoming trial. Subjects performed a tone-task similar that of Experiment 1 and classified words as either “bug,” “food,” or “tree”.

We had two reasons for adopting this particular design. First, there are several reports that, after several sessions of practice in this paradigm, subjects can sometimes learn to perform two tasks together with very little interference and no sign of a bottleneck. In fact, Hazeltine, Ruthruff, and Remington (2001) found such a result after practice using essentially the same tasks and design as the present Experiment 2. Thus, there is reason to believe that this paradigm is conducive to the overlap of central processing. Second, the use of single-task conditions seems especially likely to discourage response grouping (indeed, grouping is not even possible on the single-task trials). By deterring response grouping, we should be able to determine the true central processing modes (serial vs. parallel) adopted by a greater proportion of the subjects.

Because there was no compatibility manipulation in this experiment, we will focus only on the distribution of IRIs. The central bottleneck model again predicts that this distribution will have three specific components. We also tested a new prediction of the bottleneck model regarding the effects of SOA in the dual-task condition. The SOA manipulation we used was very modest (the tone was presented either 50 ms before or 50 ms after the word) and consequently was generally not noticed by subjects (the stimuli appear to be simultaneous). According to the bottleneck model, the task presented first should have a slight head start and therefore should tend to engage the central bottleneck first. The processing order would probably not be strictly first-come-first-served, because subjects faced with variable task orders from trial to trial have a preference to use the same central processing order from trial to trial even when the presentation order changes (e.g., De Jong, 1995). However, the slight difference in SOAs should still cause a noticeable shift in the proportions of the two processing orders. Importantly, though, the basic shape of each component of the IRI distribution should remain unchanged—the SOA manipulation should primarily cause a vertical shift in the two components corresponding to non-grouping trials (i.e., the proportion of responses in each bin should be shifted up or down by a multiplicative factor).

The predictions of the capacity-sharing model for the effects of the small SOA manipulation are less clear. However, the most straightforward prediction is that the task presented first will receive the total pool of capacity for 50 ms, before roughly equal sharing begins. The effect would be to shift the IRI distributions 50 msec to the left or the right (depending on whether the SOA was positive or negative). Thus, the IRI distributions in the -50 and +50 ms SOAs might differ primarily by a combined horizontal shift (rather than the vertical shift predicted by bottleneck models) of about 100 ms.

Method

Subjects. Twelve students from community colleges in the Mountain View, California area participated in exchange for money or partial course credit. None had participated in Experiment 1.

Stimuli. Tone stimuli were presented at 220, 880, and 3520 Hz and lasted for 300 ms. Word stimuli, subtending 1.4° horizontally by 1.9° vertically, were white against a black background. There were four bug words (ANT, FLEA, APHID, BEATLE), four food words (EGG, SOUP, CANDY, COOKIE), and four tree words (OAK, PINE, CEDAR, SPRUCE). The words remained visible until a response was made.

Procedure. Subjects responded to the 220, 880, and 3520 Hz tones by saying “low,” “medium,” or “high,” respectively. Subjects responded to bug, food, and tree words by pressing the ‘j’, ‘k’, and ‘l’ keys, respectively. Subjects were first allowed to study the words for each category. They then completed 15 blocks of 63 trials each (including three warm-up trials). The first five blocks were considered practice. Each block contained a mixture of 36 dual-task trials (18 at each of the two SOAs), 12 single-task tone trials, and 12 single-task word trials. Because subjects did not know which task(s) would appear, they presumably prepared for both of them. Subjects were asked to respond quickly and accurately to both tasks. Both tasks were given equal emphasis.

Each trial began with the presentation of a fixation cross for 500 ms, then a blank field for 300 ms, followed by the tone and/or words. If an error was made, an error message was displayed for 2000 ms. The next trial began 500 ms later. At the end of each block of trials, subjects received performance feedback (RT and percent correct) and were allowed to take a short break.

Results

Trials with either an error, an RT less than 200 ms (< 1% of all trials), or an RT greater than 3 sec (< 1% of all trials) were excluded from the analyses.

Main Effects. Mean tone-task and word-task RTs are shown in Table 2 as a function of trial type (single vs. dual) and SOA; error rates are shown in parentheses. Dual-task RTs were significantly slower than single-task RTs for both the word task, $F(1, 11) = 58.5, p < .001$, and

the tone task, $F(1, 11) = 28.4$, $p < .001$. Thus, we again found evidence of dual-task interference.

IRI Distributions. Figure 5a shows the distribution of IRIs across subjects. Plotted are the proportion of responses in each 60 ms bin; the open circles represent the -50 ms SOA condition and the filled squares represent the +50 ms SOA condition. These data clearly show the trimodal pattern predicted by the central bottleneck model. These results replicate those of Experiment 1, but with many fewer responses near zero. Thus, the data confirmed our expectations that this design (which includes a substantial number of single-task trials) would reduce the incidence of response grouping.

One of the most striking features of the IRI distributions relates to the effect of SOA. Changing the SOA appears to primarily influence the proportion of trials in the two extreme components of the IRI distribution, without noticeably changing the shapes of those distributions. The shift was in the predicted direction – e.g., when the word was presented before the tone, subjects were more likely to respond to the word first. Note, however, that the shift was not as large as one would expect from a strict first-come-first-served policy for central processing. De Jong (1995) presented evidence that central processing order tends to be partially determined in advance of the trial; specifically, subjects appear to be biased toward repeat the central processing order used on the previous trial. Indeed, when subjects responded in one order (without grouping) on a trial, they usually responded in the same order on the next trial as well (70% of the time).

To facilitate comparisons between the shape of the positive and negative components of the IRI distribution between SOAs, we replotted the IRI data in Figure 5b. Specifically, we conditionalized the data for each SOA on response order. First, we removed data from bins near zero, where response grouping might have occurred. Then, we divided the proportion of IRIs in each positive IRI bin at the 50 ms SOA by the total proportion of positive IRIs at the SOA, and we divided the proportion of IRIs in each negative bin by the total proportion of negative IRIs at that SOA. This procedure was then repeated for the -50 ms SOA as well. After making these adjustments to equalize the average height of each component at each SOA, it is easy to see that the shapes and locations of the IRI distributions are very similar for each SOA.

These results are inconsistent with the capacity-sharing model in which central operations on the two tasks share capacity roughly equally. In particular, note the sharp dip in response proportion in the bins close to zero and note the large number of responses in extreme IRI bins (beyond -450 or +450 ms). Also, as discussed earlier, the most straightforward prediction for the effect of SOA on the IRI distributions is a horizontal shift of roughly 100 ms. Contrary to this prediction, Figure 5 shows no evidence for a horizontal shift.

To provide a quantitative test between these models, we again compared the number of responses in the 90-210 ms IRI bins (both positive and negative) to the number of responses in the 270-390 ms bins (positive and negative). Even though the extreme bins are further from zero, they had roughly twice as many responses as the narrow bins (0.18 vs. 0.09); this difference was statistically significant, $F(1, 11) = 5.0$, $p < 0.05$. Eleven of the 12 subjects showed this pattern; the remaining subject appeared to have grouped responses together in close temporal synchrony (54% of responses fell between -150 and +150 ms).

Discussion

The IRI distributions obtained in Experiment 2 provide clear evidence for the three components predicted by the central bottleneck model, rather than the unimodal distribution predicted by the capacity-sharing model in which simultaneous central processes share capacity roughly equally. In addition, the effect of SOA was mainly to shift the proportion of trials within each component (i.e., a vertical shift) of the IRI distribution, without any apparent change to the position or shape of the components. This finding confirms the predictions of the central bottleneck model, but is contrary to the straightforward

prediction of the capacity-sharing model (a horizontal shift in the distributions of IRIs).

General Discussion

In this article we considered the possibility that there is graded capacity-sharing between central stages, rather than a strict (all-or-none) central bottleneck. More specifically, we evaluated the possibility that central operations can proceed in parallel, albeit more slowly than when performed in isolation. Our basic strategy was to emphasize each task equally and to encourage subjects to overlap the two tasks as much as possible.

Despite conditions that encouraged the overlap of central operations, subjects appeared to perform central operations sequentially. First, the distributions of IRIs revealed a tendency for subjects to respond to the two tasks in one of two orders or to group responses together in close temporal synchrony. This pattern was evident most clearly in Experiment 2 (see Figure 5). Thus, we replicated the basic findings of Pashler (1994a) in a design where the potential for response conflicts was minimized. This finding is consistent with the central bottleneck model, which asserts that subjects must choose one of two possible orders for sequentially performing central operations. The pattern is not consistent, however, with the model outlined in the Introduction in which subjects divide a limited pool of capacity roughly equally between central operations on the two tasks. According to that model, the IRIs should have a broad, unimodal distribution.

We performed a converging test between the candidate models (central bottleneck vs. capacity-sharing) by examining the effects of character-task S-R compatibility (Experiment 1). According to the bottleneck model, the effects of compatibility on the character task itself should be additive with the effects of SOA. In addition, these compatibility effects should carry over fully onto tone-task RT when the character-task is performed first but not when the tone task is performed first. Both of these predictions were verified in Experiment 1. According to the capacity-sharing model, however, the effects of character-task S-R compatibility on the character task itself should be larger at the 0 ms SOA than at the long SOAs. The reason is that, as noted above, the tasks must share capacity at the 0 ms SOA. The reduced capacity allocation should prolong central operations, magnifying the effects of the compatibility manipulation on the character task. However, compatibility effects did not increase at the 0 ms SOA (see Table 1); instead, they were roughly additive with SOA.

One could attempt to reconcile capacity-sharing models with the observed compatibility effects by arguing that tone-task central operations finished relatively early, so that there was little temporal overlap between central operations on the two tasks. This explanation is unsatisfactory, however, for two reasons. First, the tone-task response was often emitted after the character-task response, which suggests that there was substantial overlap. Second, if there were no temporal overlap then there should also have been no carryover of compatibility effects onto tone-task RT (see the Appendix). In other words, the combination of substantial carryover effects with no increase in compatibility effects on the character-task itself at the short SOA is incompatible with roughly equal capacity-sharing between central operations. Figure 3 shows the predictions of the capacity-sharing model under various allocations of capacity, along with the observed results (indicated by the asterisk). The observed data are consistent only with very extreme capacity allocations, such as the [0,1] and [1,0] allocations predicted by the central bottleneck model.

Revised Capacity-Sharing Models

Can the capacity-sharing model described in this paper be revised in order to account for the present data? It might be possible to account for the compatibility effects described above by postulating that more capacity is allocated to the character task on incompatible trials relative to compatible trials. For example, a subject assigned to the compatible mapping of digits onto the three response keys (1, 2, 3)

and the incompatible mapping of letters onto the three response keys (C, A, B) might allocate more capacity to the character task when a letter is presented than when a digit is presented. Compared to the original capacity-sharing model, this revised model would predict a smaller increase in compatibility effects at the short SOA, and would predict more carryover of compatibility effects on to the tone task. In other words, the predicted results would be more in line with the observed results.

Even with this modification, the revised capacity-sharing model still cannot easily account for the IRI distributions (see Figure 5), which appear to consist of three distinct components. Capacity-sharing plus random response grouping could create a spike in the center of the IRI distribution, but would not produce the dip in response proportions that was observed for IRIs near 200 ms. To explain the dip one could further propose that response grouping is not random, but occurs primarily on trials that would have produced an IRI of about 200 ms or less. This selective response grouping would take trials away from bins near 200 ms and redistribute them to bins near 0 ms. Although this revised capacity-sharing model can explain the aggregate data from Experiment 1, it is not inconsistent the observation that the subjects with the largest dips in IRIs near 200 ms showed the fewest IRIs in the bins near 0 ms, not the most (e.g., see Figure 4B). Furthermore, in Experiment 2 there were relatively few trials in any of the IRI bins between -200 and +200 ms. This pattern cannot be explained in terms of a redistribution of trials from one bin to another. To explain the IRI data, a capacity-sharing model would need to add some further provision, for example that subjects tend to use extreme capacity allocations (e.g., [.2,.8] and [.8,.2]), while rarely adopting intermediate allocations. It is not clear, however, why subjects capable of graded capacity-sharing would primarily utilize extreme allocations. In contrast, the central bottleneck model provides a clear rationale for the appearance of two distinct allocations because it states that those two allocations ([0,1] and [1,0]) are the only possible choices.

It is conceivable, of course, that subjects did share capacity between simultaneous central operations on a small proportion of trials. The present data cannot rule out this hypothesis. At the same time, there appears to be no evidence in favor of it either. The central bottleneck model with occasional response grouping can explain the present data extremely well, however, without adding the additional assumption that there was a subset of trials where capacity was shared, so parsimony would seem to favor this interpretation.

It is also conceivable that capacity-sharing can potentially occur under conditions not achieved in the present experiments. For example, it might occur after extensive practice, with greater incentives, or with much easier tasks. Further work, perhaps using the same techniques employed in the present study, is needed to evaluate these possibilities.

Structural vs. Voluntary Bottlenecks

It has been noted that the central bottleneck observed in the PRP paradigm might occur as a result of the Task-1 emphasis instructions (Meyer & Kieras, 1997a, 1997b); that is, subjects might have the latent ability to perform central operations in parallel with no interference, but instead voluntarily adopt a bottleneck strategy (e.g., to satisfy the Task-1 emphasis instructions). If so, then one might expect the central bottleneck to be eliminated when both tasks are given equal emphasis. There are now several studies, including the present one, that contradict this prediction (Carrier & Pashler, 1995; Levy & Pashler, 2001; Pashler, 1994a; Ruthruff et al., 1995, Experiment 3; Ruthruff, Pashler, & Klaassen, 2001; Tombu & Jolicoeur, 2000). It remains to be seen if the central bottleneck can be eliminated under even more extreme pressure to overlap central processing. In any case, it does not appear that the central bottleneck is caused by the instructions. These results instead support the hypothesis that – at least relatively early in practice – the central bottleneck is due to a structural limitation inherent in the cognitive architecture. In other words, the bottleneck occurs not because

subjects are encouraged to perform central operations sequentially, but rather because it is difficult to perform them in parallel.

Although it has proven difficult to eliminate the central bottleneck at relatively low practice levels, there is recent evidence suggesting that it can sometimes be eliminated after practice (Hazeltine, Ruthruff, & Remington, 2001; Hazeltine et al., 2002; Levy & Pashler, 2001; Schumacher et al., 2001). It remains to be seen whether the apparent elimination of the central bottleneck occurs because central operations are performed in parallel, or because practice in some cases eliminates the need for central operations. For instance, it is possible that there are two processing routes to activating a response code (see Lien & Proctor, 2002), one of which (response selection) comprises the central bottleneck and one of which (automatic response activation) does not. Perhaps early in practice subjects must rely on the route that does comprise the central bottleneck, but with practice they learn to rely on the processing route that does not comprise the central bottleneck (see Ruthruff, Johnston, & Van Selst, 2001).

Summary

The present dual-task experiments emphasized each task equally and encouraged subjects to overlap processing of the tasks. Substantial dual-task interference was obtained, which – a priori – could be due to a central bottleneck (i.e., serial performance of central stages) or to capacity-sharing (i.e., parallel performance of central stages, albeit with interference). The observed IRI distributions and S-R compatibility effects, however, suggest that subjects rarely, if ever, shared capacity roughly equally between tasks. Instead, it appears that subjects devoted all (or at least the vast majority) of their entire capacity first to the central operations of one task and then devoted their entire capacity to the other task. In other words, the present data support a central bottleneck model of dual-task performance. Further work is needed to determine whether the same conclusions apply to easier tasks, higher levels of practice, and greater incentives to overlap processing.

The central bottleneck model is also supported by a number of PRP studies. Because the PRP paradigm emphasizes the speed of Task 1 performance, however, those data can be explained by a voluntary (strategic) bottleneck. In contrast, the present experimental design encouraged subjects to emphasize both tasks equally and to overlap performance. Thus, the present data are consistent with the hypothesis that the central bottleneck is structural rather than voluntary.

References

- Borger, R. (1963). The refractory period and serial choice-reactions. *Quarterly Journal of Experimental Psychology*, **15**, 1-12.
- Carrier, M., & Pashler (1995). Attentional limitations in memory retrieval. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **21**, 1339-1348.
- De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 965-989.
- De Jong, R. (1995). The role of preparation in overlapping-task performance. *Quarterly Journal of Experimental Psychology*, **48A**, 2-25.
- Hazeltine, E., Ruthruff, E., & Remington, R. W. (2001). Dual-task costs depend on the pairings of stimulus and response modalities. Presented at the annual meeting of the Psychonomic Society, Orlando.
- Hazeltine, E., Teague, D., & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 527-545.
- Hommel, B. (1998). 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). *Bypassing the single-channel bottleneck in dual-task performance*. Paper presented to the 34th annual meeting of the Psychonomic Society, Washington D.C., November.
- Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology*, **36**, 138-202.
- Kahneman, D. (1983). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Keele, S. W. (1973). *Attention and human performance*. Pacific Palisades, CA: Goodyear.
- Levy, J., & Pashler, H. (2001). Is dual-task slowing instruction dependent? *Journal of Experimental Psychology: Human Perception and Performance*, **27**, 862-869.
- Lien, M.-C., & Proctor, R. W. (2000). Multiple spatial correspondence effects on dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, **26**, 1260-1280.
- Lien, M.-C., & Proctor, R. W. (2002). Stimulus-Response Compatibility and Psychological Refractory Period Effects: Implications for Response Selection. *Psychonomic Bulletin & Review*, **9**, 212-238.
- Lien, M.-C., Schweickert, R., & Proctor, R. W. (in press). Task switching and response correspondence in the psychological refractory period paradigm. *Journal of Experimental Psychology: Human Perception and Performance*.
- 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**, 1260-1280.
- 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.
- McCann, R. S., Remington, R. W. & Van Selst, M. (2000). Automaticity in visual word processing: A dual-task investigation. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 1352-1370.
- McLeod, P. (1977). Parallel processing and the psychological refractory period. *Acta Psychologica*, **41**, 381-396.
- 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 human multiple-task performance: Part 2. Accounts of psychological refractory phenomena. *Psychological Review*, **107**, 749-791.
- Meyer, D. E., Kieras, D. E., Lauber, E., Schumacher, E. H., Glass, J., Zurbriggen, E., Gmeindl, L., & Apfelblat, D. (1995). Adaptive executive control: Flexible multiple-task performance without pervasive immutable response-selection bottlenecks. *Acta Psychologica*, **90**, 163-190.
- Navon, D., & Miller, J. O. (2002). Queuing or sharing? A critical evaluation of the single-bottleneck notion. *Cognitive Psychology*, **44**, 193-251.
- 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. (1994a). Graded capacity-sharing in dual-task interference? *Journal of Experimental Psychology: Human Perception and Performance*, **20**, 330-342.
- Pashler, H. (1994b). Overlapping mental operations in serial performance with preview. *Quarterly Journal of Experimental Psychology*, **47A**, 161-191.
- Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, **41**, 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. E., & Klaassen, A. (2001). Processing bottlenecks in dual-task performance: Structural limitation or voluntary postponement? *Psychonomic Bulletin & Review*, **8**, 73-80.
- Ruthruff, E., Johnston, J. C., & Van Selst, M. A. (2001). Why practice reduces dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, **27**, 3-21.
- Ruthruff, E., Johnston, J. C., Van Selst, M. A., Whitsell, S., & Remington, R. (in press). Vanishing dual-task interference after practice: Has the bottleneck been eliminated or is it merely latent? *Journal of Experimental Psychology: Human Perception and Performance*.
- Schumacher, E. H., Seymour, T. L., Glass, J. M., Kieras, D. E., & Meyer, D. E. (2001). Virtually perfect time sharing in dual-task performance: Uncorking the central attentional bottleneck. *Psychological Science*, **121**, 101-108.
- Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology*, **18**, 105-139.
- Schweickert, R., & Townsend, J. T. (1989). A trichotomy: Interactions of factors prolonging sequential and concurrent mental processes in stochastic discrete mental (PERT) networks. *Journal of Mathematical Psychology*, **33**, 328-347.
- Sommer, W., Leuthold, H., & Schubert, T. (2001). Multiple bottlenecks in information processing? An electrophysiological examination. *Psychonomic Bulletin & Review*, **8**, 81-88.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica*, **30**, 276-315.
- Tomblu, M., & Jolicoeur, P. (2000). Is the PRP effect due to a strategic or structural bottleneck? Presented at the annual meeting of the Psychonomic Society, New Orleans.
- Tomblu, M., & Jolicoeur, P. (in press). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*.
- 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.

Appendix

Capacity-Sharing Model Predictions for the Effect of S-R Compatibility

In this Appendix we determine how manipulations of the duration of central processing should affect dual-task RTs, according to the capacity-sharing model outlined in the Introduction (where central stages share capacity, but input and output stages do not). Specifically, we want to determine what effect a central stage manipulation on Task Y will have on Task Y itself and how much of this effect will carry over onto Task Z. We focus on the central stage manipulation used in Experiment 1: character-task S-R compatibility. It is convenient to model the effect of incompatibility as adding a central stage (stage "I") beyond that required in the compatible condition. The compatibility effect on the character task itself, therefore, is simply the average time to complete stage I. Let C_Y and C_Z represent the proportion of capacity (assumed to be constant across trials) allocated to Task Y and Task Z when their central operations overlap in time, where $C_Y + C_Z = 1$. Let W represent the amount of work required on a particular central stage, defined as the time (in milliseconds) it would take to complete that stage under full capacity. Under reduced capacity, the work accomplished in time T is proportional to the capacity allocated to that task; i.e., $W = T * C$.

Compatibility Effects on Task Y

The compatibility effect on Task Y at the long SOA ($CE_{Long\ SOA}$), where Task Y is assumed to operate at full capacity, is simply the average time required to complete stage I. Thus, it is equal to the expected value of W_I .

$$CE_{Long_SOA} = E [W_I]$$

At short SOAs, however, stage I will sometimes share capacity with Task Z; the reduced capacity will prolong stage I and hence increase CE_{short_SOA} relative to CE_{long_SOA} . Let $T_{Overlap}$ be a random variable representing the amount of time during which stage I of Task Y overlaps with central operations on Task Z (i.e., the amount of time stage I operates with reduced capacity). Let $W_{Overlap,Y}$ represent the amount of work accomplished on Task Y during $T_{Overlap}$.

$$W_{Overlap,Y} = T_{Overlap} * C_Y. \quad [1]$$

CE_{short_SOA} , the average time to complete stage I at the short SOA, is equal to the expected value of $T_{Overlap}$ plus the expected value of the time to complete the remaining work ($W_I - W_{Overlap,Y}$) at full capacity. Therefore, $CE_{short_SOA} = E [T_{Overlap}] + E [W_I] - E [W_{Overlap,Y}]$. Substituting the value of $W_{Overlap,Y}$ from Equation 1, and noting that $C_Z = 1 - C_Y$, it follows that

$$CE_{short_SOA} = E [W_I] + E [T_{Overlap}] * C_Z.$$

The average increase in the compatibility effect at the short SOA relative to the long SOA (denoted by ΔCE) is equal to CE_{short_SOA} minus CE_{Long_SOA} . Thus,

$$\Delta CE = E [W_I] + E [T_{Overlap}] * C_Z - E [W_I] \quad CE_{Long_SOA}$$

$$\Delta CE = E [T_{Overlap}] * C_Z \quad [2]$$

As an informal explanation of Equation 2, note that $T_{Overlap} * C_Z$ is the amount of work completed on Task Z while it overlaps with stage I. This work on Task Z means lost work on Task Y, which must be made up later at a cost of one millisecond per each unit of lost work. The maximum possible ΔCE is obtained with maximum $T_{Overlap}$, which occurs when all the work on stage I of Task Y is performed under reduced capacity. Thus, $\text{Max}(T_{Overlap}) = W_I / C_Y$. Under this condition, ΔCE will be equal to $CE_{Long_SOA} * (C_Z / C_Y)$. With equal capacity sharing ([0.5,0.5]), for example, ΔCE would be equal CE_{Long_SOA} ; thus, CE_{short_SOA} would be twice as large as CE_{long_SOA} .

Carryover of Compatibility Effects onto Task Z

The carryover of the Task Y compatibility effect onto Task Z is simply the lengthening of Task Z central operations during the overlap with stage I of Task Y. Let $W_{Overlap,Z}$ represent the amount of work accomplished on Task Z during $T_{Overlap}$ at reduced capacity C_Z on a given trial.

$$W_{Overlap,Z} = T_{Overlap} * C_Z. \quad [3]$$

Table 1

Response time and error rate (in parentheses) for each task as a function of the SOA and character-task stimulus-response mapping (compatible vs. incompatible) in Experiment 1.

	SOA (ms)		
	-1000 (Char then Tone)	0 (Simultaneous)	+1000 (Tone then Char)
Tone Task			
Compatible	759	1051	891
Incompatible	793	1118	869
Character Task			
Compatible	740 (0.2)	996 (0.7)	668 (0.1)
Incompatible	900 (4.3)	1157 (3.9)	869 (6.6)

The time cost (on Task Z) of the overlap is the time required to complete $W_{Overlap,Z}$ under reduced capacity (i.e., $T_{Overlap}$) minus the time it would have taken to complete the same amount of work under full capacity. Thus, the time cost is equal to $T_{Overlap} - W_{Overlap,Z}$. Using Equation 3, this time cost can also be expressed as $T_{Overlap} - T_{Overlap} * C_Z$ which reduces to $T_{Overlap} * C_Y$. Thus, the average carryover effect is given by

$$\text{Carryover} = E [T_{Overlap}] * C_Y. \quad [4]$$

As an intuitive explanation, note that $T_{Overlap} * C_Y$ is the amount of Task Z work lost to Task Y during the overlap, to be made up later at a cost of one millisecond per unit of lost work.

The maximum carryover effect (obtained with maximum $T_{Overlap}$) is equal to $E [W_I]$, which is simply the baseline (i.e., long SOA) compatibility effect on Task Y (i.e., CE_{Long_SOA}). Thus, the baseline compatibility effect can carry over fully onto task Z, but no more.

Carryover vs. ΔCE

Note that ΔCE (see Equation 2) and carryover (see Equation 4) are linearly related to each other: $\Delta CE = \text{Carryover} * (C_Z / C_Y)$. Also note that the maximum possible carryover and maximum possible ΔCE are both proportional to the baseline compatibility effect (CE_{Long_SOA}). Therefore, it is convenient to divide each term by CE_{Long_SOA} , so that their maximum values do not depend on CE_{Long_SOA} .

$$\frac{\Delta CE}{CE_{Long_SOA}} = \frac{\text{Carryover}}{C_Y} * \frac{C_Z}{CE_{Long_SOA}} \quad [5]$$

This linear relationship is shown in Figure 3 for several different values of C_Y and C_Z . It should not be surprising that these two measures are tightly coupled. Informally speaking, the effects of compatibility on Task Y itself and the carryover onto Task Z both depend primarily on the amount of work lost to the other task during the temporal overlap of their central operations.

Acknowledgments

This research was supported by grants from the National Research Council and the National Aeronautics and Space Administration. The research was also supported by grant MH45584 from the National Institute of Mental Health to Harold Pashler. We thank James Johnston, David Meyer, Mei-Ching Lien, and Roger Remington for helpful comments. Please send correspondence to Eric Ruthruff, NASA – Ames Research Center, Moffett Field, CA 94035, or eruthruff@mail.arc.nasa.gov.

Table 2

Response time (in milliseconds) and error rate (shown in parentheses) for each task as a function of the number of tasks (single, dual) and the SOA (-50 ms, +50 ms) in Experiment 2.

	Single	Dual	
		Word-Tone -50 ms	Tone-Word +50 ms
Tone Task	722 (5.5)	905 (6.8)	871 (7.4)
Word Task	755 (4.8)	1013 (6.5)	1042 (6.3)

Figure Captions

Figure 1: Example processing time diagrams for a central bottleneck model of Experiment 1. The left panels correspond to trials where tone-task central operations are performed first; the right panels correspond to trials where the character-task central operations are performed first. The top panels correspond to the compatible stimulus-response mapping on the character task; the bottom panels correspond to the incompatible mapping.

Figure 2: (A) Example processing time diagram for the capacity-sharing model. (B) Central operations on Tasks Y and Z share a limited pool of capacity over time.

Figure 3: Capacity-sharing model predictions for the short SOA. Shown are the predicted increase in the effects of the compatibility manipulation on the character task itself at the short SOA (relative to the long SOA) and the predicted carryover of compatibility effects onto the tone task. Predictions are shown for several possible allocations of capacity to the two tasks: [0,1], [.67,.33], [.5,.5], [.33,.67], and [1,0], where the first and second numbers in brackets indicate the capacity initially allocated to the character and tone tasks, respectively. The asterisk represents the results from Experiment 1 (with standard error bars).

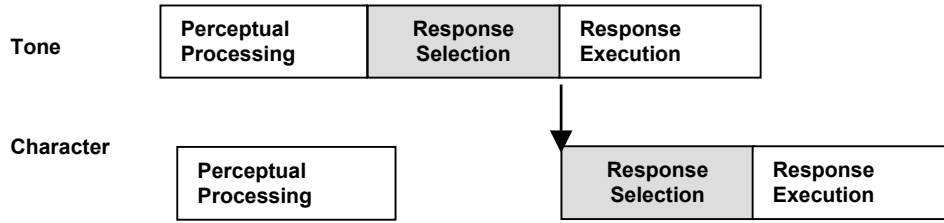
Figure 4: Experiment 1 - Proportion of inter-response-intervals (IRIs) in each 60 ms-wide bin. Positive IRIs mean the subject responded to the tone task before the character task. (A) Compatible (open squares) versus incompatible (filled circles). (B) Majority Group (few IRIs near zero; diamonds) versus minority group (many IRIs near zero; plus sign).

Figure 5: Experiment 2 - Proportion of inter-response intervals (IRIs) in 60 ms-wide bins as a function of SOA (-50 ms vs. +50 ms, where positive SOAs indicate that the tone appeared before the character). Positive IRIs mean the subject responded to the tone task before the character task. (A) Actual data; (B) Data adjusted to equalize area under each curve.

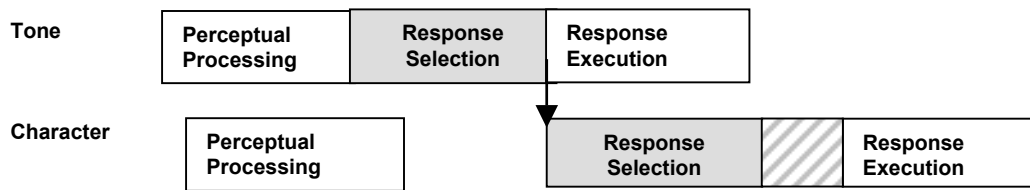
Figure 1

Tone Before Character

Compatible Character Task

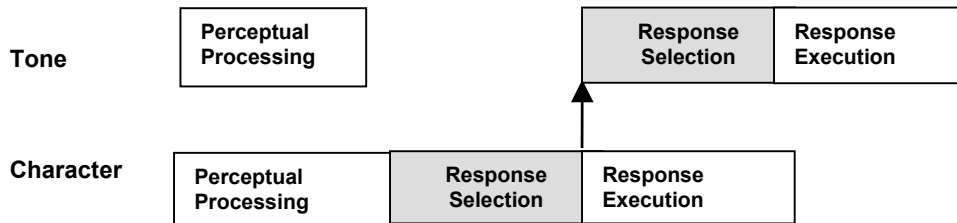


Incompatible Character Task



Character Before Tone

Compatible Character Task



Incompatible Character Task

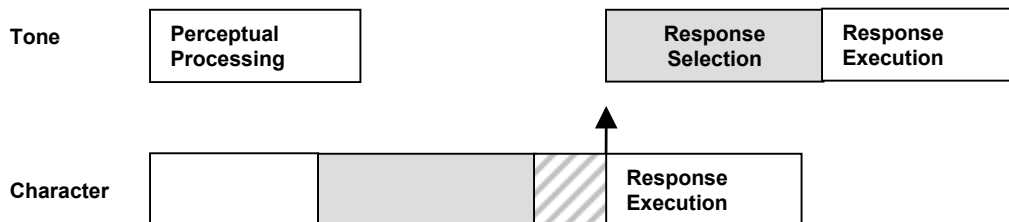
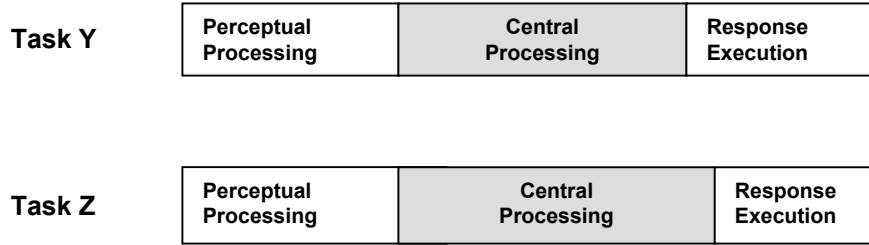


Figure 2

A



B

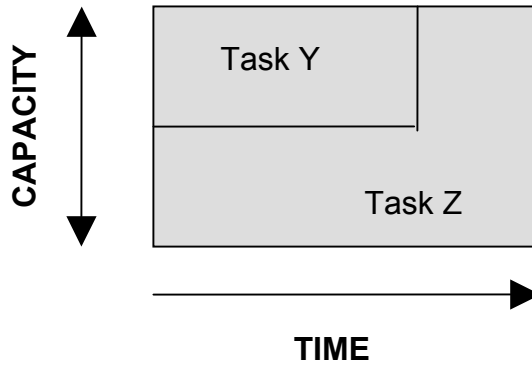


Figure 3

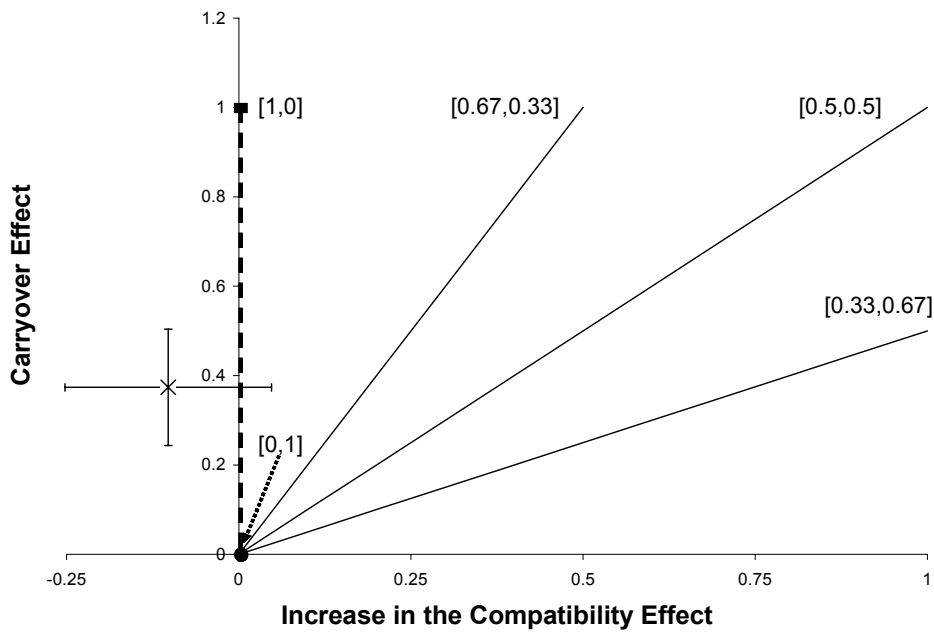


Figure 4

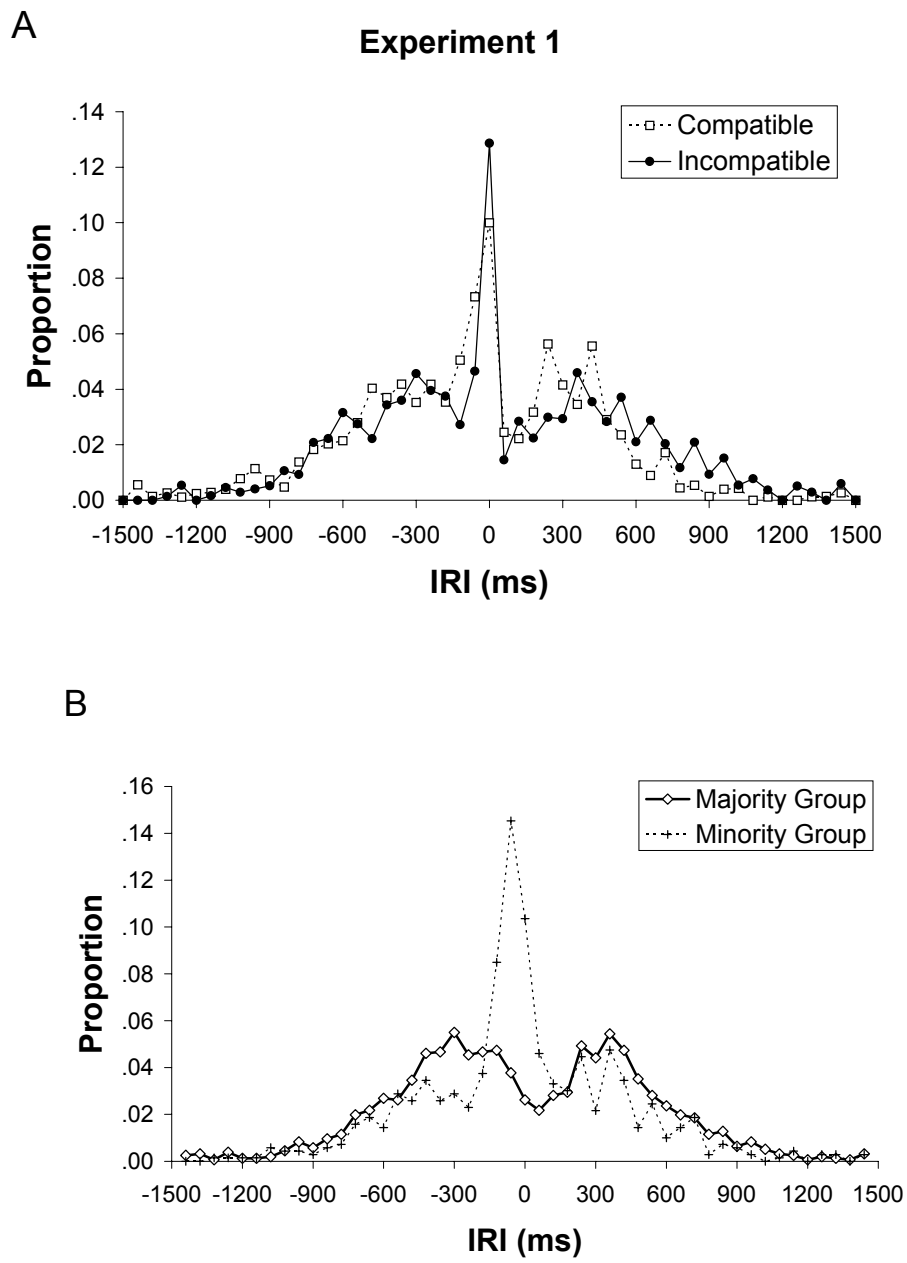
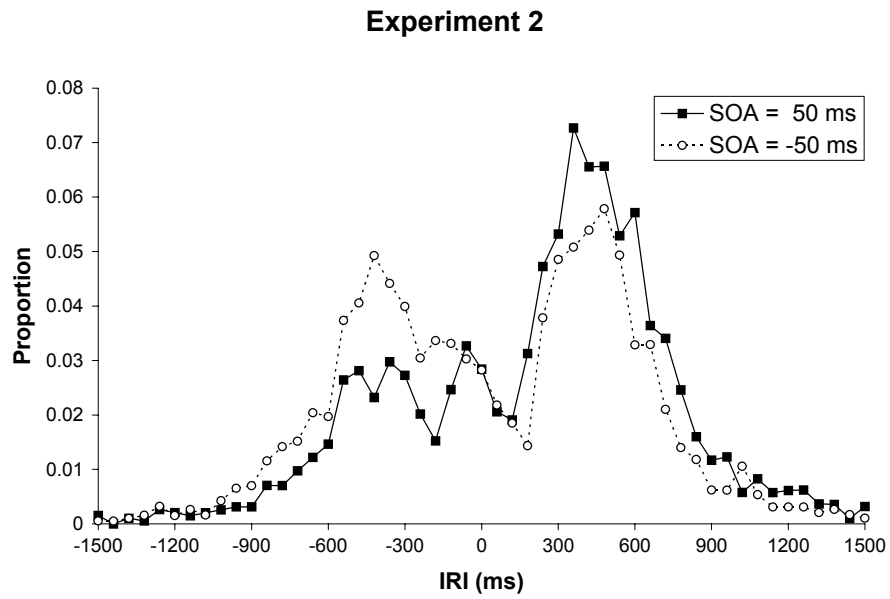
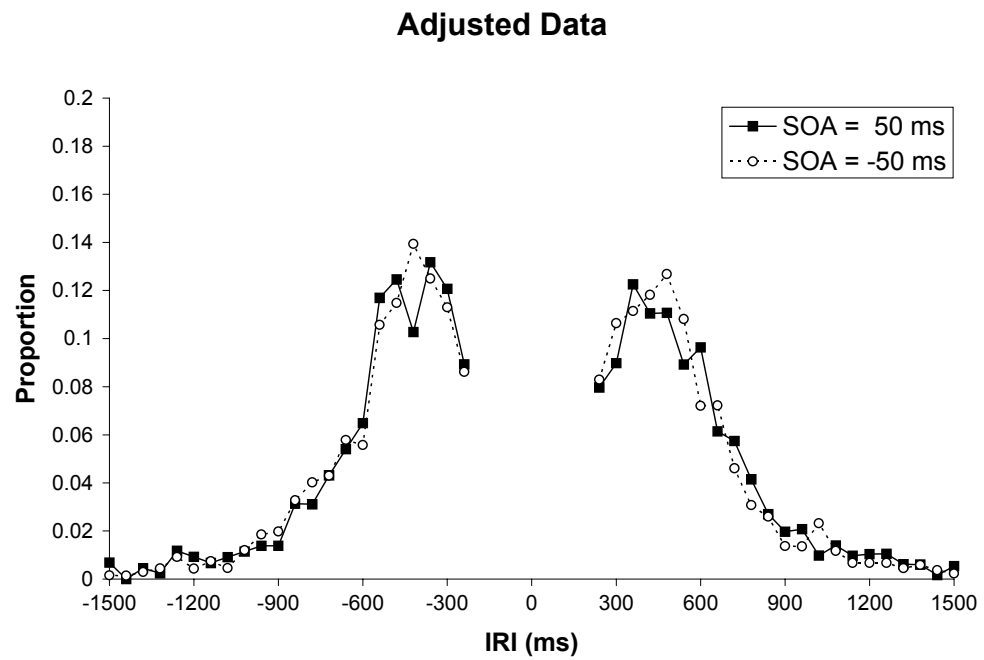


Figure 5

A



B



Footnotes

¹ Note that individual subjects might show only one or two of the predicted components.

² The reallocation assumption is needed to explain the PRP effect. If capacity were not reallocated, then task-2 RT would be slow at all SOAs, not just the short ones.

³ Rather than assuming an instantaneous capacity shift, one could allow for a switch cost. The inclusion of a modest switch time would not appear to alter the predictions discussed below.

⁴ A complication arises if one considers the possibility that capacity allocation varies considerably from trial to trial. Nevertheless, each point on Figure 3 is consistent only with certain combinations of capacity allocations. For instance, a finding of zero increase in compatibility effects combined with substantial carryover (as we found in Experiment 1) suggests that intermediate capacity allocations did not occur.

⁵ The predictions shown in Figure 3 are expressed as a ratio of the baseline (long SOA) compatibility effect, so that they will be independent of the effect size observed in a particular experiment. To place Experiment 1 on this plot, it was necessary to rescale the values reported above. The mean character-task compatibility effect was 161 ms at the short SOA and was 180.5 at long SOAs (-1000 and +1000 ms). Thus, the compatibility effect at short SOAs was 0.89 times that found at the long SOAs, producing an “increase” of -0.11. The carryover of S-R compatibility onto the tone task was 67 ms, which is 0.37 times the baseline (long SOA) compatibility effect on the character-task.