

Queuing or Sharing? A Critical Evaluation of the Single-Bottleneck Notion

David Navon and Jeff Miller

The University of Otago

The model of a single central bottleneck for human information processing is critically examined. Most evidence cited in support of the model has been observed within the overlapping tasks paradigm. It is shown here that most findings obtained within that paradigm and that were used to support the model are also consistent with a simple resource model. The most prominent findings are the millisecond-for-millisecond slope at the left of the RT₂–SOA curve, the high RT₁–RT₂ correlation, the additivity of the effects on RT₂ of SOA and of the difficulty of selecting R₂, and the washout of the effect of S₂ discriminability on RT₂ in a dual-task condition. In addition, the asymmetry of the effects of the dual-task requirement on RT₁ and RT₂ can be accounted for by the resource model provided that it assumes uneven allocation of resources, which is quite reasonable in view of the task asymmetry inherent in the demand characteristics of the paradigm. The same is true for two other findings that appear to support the single-bottleneck model—that in the dual-task condition, the demand of the first task affects equally RT₁ and RT₂ and that its effect on RT₁ is the same as the corresponding effect in the single-task condition. Furthermore, the single-bottleneck model is hard to reconcile with a negative slope at the left of the RT₁–SOA curve or a positive slope at the left of the IRI–SOA curve, unless augmented by ancillary assumptions that are yet to be substantiated. Representative data were fit by each of the models using its optimal set of parameters. Both models achieved quite good degrees of fit. It is further argued that since the overlapping tasks paradigm is heavily biased in favor of a speedy reaction to the stimulus that appears first, it is nonoptimal for testing the central bottleneck model. Finally, the bottleneck model is examined in terms of other scientific criteria. © 2001 Elsevier Science (USA)

Key Words: single-channel hypothesis; central bottleneck; single-bottleneck model; psychological refractory period; PRP; overlapping tasks paradigm; response grouping; task preparation; resource theory; outcome conflict; dual-task performance; dual-task deficit; strategic queuing.

Modeling human information processing seems at times discouragingly complex. No wonder that students of cognitive psychology welcome any

This article was conceived in 1996 while the first author was a visiting professor at the University of Otago. We are indebted to Claus Bundeson, John Duncan, Gordon Logan, and Harold Pashler for helpful comments made on previous versions of this paper.

Address correspondence and reprint requests to David Navon, Department of Psychology, University of Haifa, Haifa 31905, Israel. E-mail: dnavon@psy.haifa.ac.il.



sign of possible simplicity. Theorizing about dual-task interference seems to have been affected by such a tendency.

Of all the ways by which dual-task interference has been explained (see reviews in Gopher & Donchin, 1986; Meyer & Kieras, 1997a; Navon & Gopher, 1979; Navon & Miller, 1987; Pashler, 1998; Wickens, 1984), one of the most appealing is the notion of a single serial bottleneck, namely a structure limited to processing only one thing at a time. This account is most appealing because it appears to be simple, economical, and congruent with phenomenal experience. Ironically, however, the appeal entails some risk: That we are tempted to accept the account without paying due attention to the weight of evidence that selectively supports it. The temptation has long been with us (e.g., Bertelson, 1966; Borger, 1963; Craik, 1947, 1948; Creamer, 1963; Davis, 1957; Keele, 1973; Smith, 1967; Telford, 1931; Welford, 1952, 1967; see reviews in Kinsbourne, 1981; Pashler, 1984, 1994a). Recently it seems to have been growing (McCann & Johnston, 1992; Pashler, 1993, 1994a, 1998). Whereas quite a bit of evidence has been marshaled to substantiate the notion of a single bottleneck, it has not always been checked whether that evidence could not be reconciled with alternative accounts or even whether the paradigm generating that evidence is most suitable for establishing the existence of a bottleneck structure. In this article we examine both of these issues.

THE SINGLE-BOTTLENECK NOTION AND ITS RIVALS

Is a Single Bottleneck Needed to Explain Queuing of Tasks?

A single bottleneck would force tasks to queue for its services. Accordingly, results that seem compatible with queuing have been used to support the single-bottleneck notion.

It would be quite perplexing if instances of queuing were never observed in a system as complex as the human information processing system must be. The issue, however, is how sweeping should be the generalization from what appear to be demonstrations of such queuing. Proponents of the original single-channel hypothesis (e.g., Welford, 1952, 1967) had a fairly clear stance. They seem to have meant that notion as a grand unifying principle for human performance. Whoever claimed at the time to have disproved it (e.g., Allport, Antonis, & Reynolds, 1972) argued that it did not stand that pretense, not that processes never queue for any structure. More recent versions of the hypothesis are certainly more theoretically circumscribed, but still aim at quite a high target: Showing that most of the variance in task interference is due to queuing for a single central processing structure (e.g., Pashler, 1993).

That target might appear somewhat ambitious, considering that a similar target has apparently been missed by resource theory (e.g., Navon, 1984).

On just an a priori basis, how likely is it that a specific hypothesis, as the modern formulation of the single-bottleneck notion is, could succeed where more flexible models assuming a single pool of resources, let alone multiple-resource models, failed?

Note that the single-bottleneck notion and resource theory are neither adversaries nor independent. A resource might be weakly defined as any provision for processing (e.g., processing device, storage units, communication device, and energy). In that sense, resource theory actually subsumes the single-bottleneck notion, where a serial bottleneck is regarded as a special type of resource—a unipartite commodity that can be used by different processes only in succession—in contrast with resources that are divisible.¹ The single-bottleneck notion, much like divisible resource models, hinges on the validity of a postulate that it shares with those models, namely that task interference results from scarcity of some provision for processing. Validating that postulate requires ruling out that the particular interference to be accounted for is actually due to cross-talk or other conflicts between ongoing processes, a fairly tough enterprise (Navon, 1984). Furthermore, the single-bottleneck notion assumes that the scarce provision is not only indivisible but also required by all controlled processes. In that it resembles one prominent instance of the possible set of resource models, namely the strict central

¹ Resource theorists (e.g., Broadbent, 1971; Gopher, 1986, 1993, 1994; Kahneman, 1973; Kantowitz & Knight, 1976a, 1976b; McLeod, 1977; Moray, 1967; Navon, 1984; Navon & Gopher, 1979, 1980; Norman & Bobrow, 1975; Sanders, 1979; Wickens, 1984) have been quite open-minded regarding the source of the limit on concurrent processing which they have called "resource." Broadbent, for example, presented the quite general notion of "an information processing channel with limited capacity." Other resource theorists were not much more restrictive. They typically regarded resources as divisible, but some of them plainly stated that division might be done by time sharing a unipartite structure. For example, Norman and Bobrow explicated the notion of mental resources by drawing the analogy to resources used by a digital computer, including what they called processing effort, namely CPU time (1975, p. 45). Navon and Gopher specifically disclaimed addressing the temporal dimension of processing, since they believed it was hard to decide empirically (1979, p. 249). Navon (1984) followed that tradition by defining the term resource "with no reference to the overall number of input units" (1984, p. 217). It thus seems that the major conceptual distinction is defined not just by number of input units but also by mode of service: On the one hand, a unipartite device restricted to strict queuing of entire processes, and on the other hand, provisions that can afford service to concurrent processes in some manner, be it parallel processing or rapid switching of a single device. As long as there is no way to distinguish empirically rapid switching from parallel processing, both must be classed in the same category. Should we refer to that category by the term "divisible resources" (and use "indivisible resource" to refer to the former category), or should we rather restrict the term "resources" just to the latter (and use the term "bottleneck" to refer to the former)? One of us has previously suggested that "resources" be qualified by five properties (aggregate nature, exclusive usage of units, distributability, effectiveness, and scarcity) and that provisions that do not satisfy these properties be called "commodities" (Navon, 1985). Accordingly, we reserve below the term "resource" to provisions that can be divided in either way, but use "divisible resources" when that is required to avoid confusion.

capacity model. The latter, however, fails to accommodate a number of phenomena of task interference (Navon & Gopher, 1979; Sanders, 1979; Wickens, 1980, 1984).

These problems aside, there is still room for challenging an assumption that divisible resource models do not share with the single-bottleneck notion, namely divisibility. Unfortunately, that is not much simpler.

It might be thought that the divisibility assumption is rendered implausible by the frequent occurrence of queuing. Note, however, that although indivisibility entails queuing, evidence for queuing does not rule out divisibility. It demonstrates, of course, that parallel processing with unlimited capacity and no interprocess interaction is not always possible—which nobody ever denied—and that a serial strategy is often preferable—which everybody would be reluctant to challenge (see Townsend & Ashby, 1983). Serial processing of tasks must be quite prevalent, especially when the setting appears to induce it. After all, if concurrent processing of tasks (namely time-sharing or parallel processing) were possible, it presumably would require a surge of effort (Kahneman, 1973; ch. 2) that may incur some extra cost (Navon & Gopher, 1979, p. 229) or feel aversive (Navon, 1989a, 1989b), it would depend heavily on practice (Hirst, Spelke, Reaves, Caharack, & Neisser, 1980; Logan, 1979, 1988; Schneider & Shiffrin, 1977), it could produce motor or other sorts of conflict (Hirst & Kalmar, 1987; Logan & Schulkind, 2000; Meyer & Kieras, 1997b; Navon & Miller, 1987), and/or it would be susceptible to errors (Duncan, 1980). Serial processing might be a natural strategy for overcoming these problems (cf. Kinsbourne, 1981; Meyer & Kieras, 1997b), often the optimal strategy (see Logan & Gordon, 2001). Accordingly, divisible resource models need not—and typically did not—claim that concurrent processing was the rule, only that it was possible.

Another false belief about divisible resource models is that they entail full concurrence. Since concurrence and queuing are conceived of as opposite categories, anything that might suggest queuing is taken as evidence against resource models. However, does indeed the notion of resource division entail full concurrence? First, even if resource division was possible, queuing could be strategically opted for by allocating all resources first to one task, then to the other one. Furthermore, even when resources are allocated by demand, full concurrence may not be attained. If all resource units in a pool were available at any moment for reallocation with no cost, no queuing would be required. That, however, is not always the case. When one stimulus arrives slightly before the other one, and some resources are already engaged by it when the second one arrives, it may be inefficient to reallocate some of them abruptly to the latter. In such a case, resource allocation could in theory produce a mixture of concurrence and queuing.²

² Consider, for example, a hotel desk that controls nine bellboys, each capable of carrying a single suitcase. Suppose two hotel residents, each having six suitcases, call the desk with

Now, could the notion of a single bottleneck realistically be expected to predict the extent of dual-task interference? It would have, had the bottleneck applied to the entire course of processing. That, however, is clearly not the case. Most often there are some savings from performing tasks concurrently (cf. reviews in Kahneman, 1973; Kerr, 1973; Moray, 1967; Wickens, 1984): The time to perform two concurrent tasks is typically shorter than the sum of the times to perform each separately. Hence, the bottleneck, if it existed, would not be the processing system in toto, but rather a specific structure within it that is probably brought to bear at a specific stage (Pashler, 1984, 1994a). The saving due to concurrence must thus be attributed to other stages that are possibly parallel. Hopefully, the bottleneck would apply to a given, specifiable stage of processing.

A number of studies have localized the bottleneck in the stage of response selection (e.g., Fagot & Pashler, 1992; McCann & Johnston, 1992; Pashler & Johnston, 1989; see review in Pashler, 1994a), though some scholars insist that the locus is postselectional (e.g., Keele, 1973; Meyer & Kieras, 1997b; Shumacher et al., 1999), some others allow for both loci (e.g., de Jong, 1993; Ehrenstein, Schweikert, Choi, & Proctor, 1997; Logan & Burkell, 1986) and still others refine the localization, arguing that it is not in S-R translation but rather in the central control of response activation (Hommel, 1998). However, even if the locus were known with certainty, that would not let us predict readily the amount of task interference, since stage durations depend on a host of factors.

What Actually Is the Main Issue?

How strong is the evidence in favor of the single-bottleneck notion? It depends, of course, on what is taken to be its rival. We chose to pit it against the notion of divisible resources.

The current tide goes against that notion. Indeed, the existence of processing resources has too long been taken for granted. Whoever worried about that (e.g., Allport, 1980; Navon, 1984, 1985) presumably did not mean that it was time for the pendulum to swing in the other direction, leading people to dismiss divisible resource models just on the grounds that they were overly variegated. Resources do exist in nature, and their effects are variegated. Whether they exist in our information processing system is an empirical question that is still worth serious consideration. In addition, the notion of divisible resources has been formalized as a tractable theory assum-

a small time lag, and each asks for luggage assistance. If six bellboys were already carrying the luggage of the first customer when the second resident called, the desk manager would probably let them complete their missions to avoid reallocation cost. Hence, three suitcases of the second resident would be carried immediately in concurrence with the suitcases of the first resident, whereas the remaining three would queue for service. Thus, resource allocation could produce a mixture of concurrence and queuing, unless resource units were in themselves infinitely subdivisible. That queuing would, however, be partial rather than strict.

ing both concurrence and processing limits. Since at present there is no other comparably formal model of concurrent processing with limit, resource theory serves at least as the best known exemplar of a class of possible models that posit concurrent processing and can accommodate interference (e.g., models of outcome conflict; see Navon, 1985; Navon & Miller, 1987).

Thus, in this article we examine mainly evidence meant to test between models of dual-task performance assuming a single bottleneck and models assuming divisible resources. The question we address is thus not whether task interference is temporally localized at a certain processing stage or where the locus is, but rather what the source of the interference is.

The trouble is that both notions—that of a single bottleneck on the one hand and that of divisible resources on the other hand—are sufficiently fuzzy to allow quite a bit of post hoc polemic. Hence, before we move on to examine the evidence, it might be useful to reformulate the issue, list all logical options, and try to map existing models onto that list.

Three questions may be asked about queuing in dual-task performance: First, is there any point during cognitive processing at which queuing is *universally* forced by some structural limit, such as a single bottleneck? Second, if there is not, is queuing *sometimes* forced by some structural limit? Third, since queuing might be due to reasons other than a structural limit (strategic choice, for one), how prevalent is queuing anyway? The possible answers to these questions fall into seven major categories: (1) queuing is *always* forced; (2) queuing is *sometimes* forced, but when it is not forced, it is *always* used anyway; (3) queuing is *sometimes* forced, but when it is not forced, it is *sometimes* used anyway; (4) queuing is *sometimes* forced, and when it is not forced, it is *never* used; (5) queuing is *never* forced, but it is *always* used anyway; (6) queuing is *never* forced, but it is *sometimes* used anyway; and (7) queuing is *never* forced, furthermore, it is *never* used.

What are the stances of the two major rival theories with respect to these possibilities? To avoid possible misattribution, let us just map both theories into broad unions of possibilities. It seems that any model of a single bottleneck must fall somewhere within the union of possibilities 1, 2, and 5, where the notion of a structurally limited bottleneck corresponds to possibility 1. In contrast, any model of divisible resources—and for that matter any other model that allowed concurrent processing—must fall somewhere within the union of possibilities 3 and 6. Possibilities 4 and 7, entailing that nonforced queuing is practically a nonexistent phenomenon, were never seriously proposed as far as we know. The reason is that if resources were divisible, then any division would be feasible in principle, including a 100%–0% allocation which might be preferable in many cases.

The main issue, then, is the extent to which experimental evidence can discriminate between these logical alternatives.

The Overlapping Tasks Paradigm

Most of the data cited in support of single-bottleneck models are to be found within what is often referred to as the overlapping tasks paradigm³ (Bertelson, 1967; Borger, 1963; Carrier & Pashler, 1995; Creamer, 1963; Herman & Kantowitz, 1970; Johnston, McCann, & Remington, 1995; Kantowitz, 1974; Karlin & Kastenbaum, 1968; McCann & Johnston, 1992; Pashler, 1984; Pashler & Johnston, 1989; Welford, 1967; see reviews in Kinsbourne, 1981; Pashler, 1994a; Pashler & Johnston, 1998; van Selst, Ruthruff, & Johnston, 1999): The stimuli for two speeded choice tasks are presented with a positive SOA, namely the presentation of the stimuli relevant for task 2, S2, is delayed with respect to presentation of stimuli relevant for task 1, S1. Reaction times for both task 1, RT1, and task 2, RT2, are measured.

The paradigm has been very productive. Numerous experiments have used it to study various aspects of dual-task performance. Several robust findings in it seem quite compatible with the notion of a single serial bottleneck that is probably central (see reviews in Pashler, 1994a, 1998; Pashler & Johnston, 1998). This was nailed down by explicating the predictions from a formal model that attributes task interference to postponement of central processing by such a bottleneck (Pashler, 1984, 1994a). Let us refer hereafter to this model by the term *Single-Bottleneck Model*.

It can hardly be disputed that the paradigm is well suited for studying the locus of a limit—whatever that limit might be. We are, however, concerned with another question—how suitable is it for studying the nature of that limit.

First, note that the paradigm is in a way biased in favor of queuing by virtue of stressing the speed of responding to two discrete, simple tasks. It is shown below [Appendix B, Eq. (6); see also “Prediction 1: RT2 Depends Heavily on SOA”] that provided that switching time is negligible, queuing minimizes RT1 yet leaves RT2 almost as it would have been had resources been divided between tasks. Since mean RT across tasks is thus minimized, queuing helps to accomplish an invited objective (cf. Logan & Gordon, 2001).

Second, due to the positive SOA dictated by the paradigm, task 1 is most often being processed by the time task 2 is ready to be processed. Hence,

³ The paradigm is also often called the PRP (psychological refractory period) paradigm. The choice of this term seems unfortunate because it presupposes a particular explanation for its findings. Worse yet, it presupposes an account that is somewhat old-fashioned. The term connotes the idea of temporary paralysis following processing or at least considerable discreteness of processing. Early users of this term (Craik, 1948; Telford, 1931) seem to have really meant the latter idea. Welford (1959, 1967) adopted it as one possibility. However, modern accounts (e.g., Kinsbourne, 1981; Pashler, 1994a) do not resort to it at all.

resource division requires reallocation. The cost of reallocation might offset any advantage that resource division has when $SOA = 0$, like the fact that it does not incur a switch cost, which queuing does. That must increase the relative advantage of queuing.

Furthermore, the special demands of the paradigm are highly asymmetrical with respect to the two tasks. In most cases the instructions squarely ask subjects to respond as rapidly as possible to the first stimulus. (For example, a not uncommon phrasing is “. . . respond quickly and accurately to both tasks, but give special emphasis to responding quickly to task 1. Try to respond immediately to the task 1 stimulus.”) The bias in the modal instructions is not moderated by any incentive for concurrent processing, let alone for equal emphasis. Despite the mild warning to take both tasks seriously, subjects are liable to feel invited to focus exclusively on task 1 before they move on to task 2. The call for heavy emphasis on task 1 exacerbates the bias that already exists in the paradigm: The fact that the tasks are always presented with a positive SOA, and that the subject is told that the stimulus for one of the tasks would always come first, could in itself induce subjects to process the tasks serially. As argued above, subjects may do that for a variety of reasons even when there is no special incentive. But in the special circumstances of the paradigm where S1 always comes first, so that it makes sense to start processing it as soon as it is presented, subjects might tend to finish processing task 1 and “get it over with.” Such a strategic opting for queuing does not indicate that the option of dividing resources between tasks is infeasible in any other circumstances (cf. Pashler, 1994a, p. 230). Actually, when the instructions are neutral, RT1 is more sensitive to SOA (Pashler, 1991, Exp. 3), and when subjects are not informed of the order of stimuli (though they might learn it), the effect of SOA on RT1 is even higher (Pashler, 1990).

In view of the bias in favor of queuing built into the paradigm, evidence of queuing observed within it is not incompatible with resource theory, even in its stronger version (namely possibility 6 in the list above): Subjects could in theory allocate all their resources to task 1, then shift all of them to task 2. More important, even if subjects opted for some concurrence, their allocation of resources would presumably be heavily biased in favor of task 1. Note that, whereas evidence for strict queuing does not rule out resource theory, evidence for concurrence, uneven as it may be, would be strong evidence against the bottleneck notion and would lend some support to resource theory. Accordingly, we first examine whether the results obtained within the paradigm unquestionably indicate queuing in the first place.

Answering that in the affirmative requires more than a good fit with a model of strict queuing. Due to the inherent bias of the paradigm, it would have been surprising if such a fit had not been found. The question is whether the same data could not be accommodated equally well by a model that allows concurrence. Some results might be compatible with *any* degree of

concurrence, still others with a limited degree of it. Since even emphasis on both tasks is a priori unlikely in view of the demand characteristics of the paradigm, a model that posits concurrence with more emphasis on task 1 is a reasonable alternative to be pitted against the model of strict queuing.

So, how does the single-bottleneck notion fare in comparison with alternative models in its own court? Before examining that, let us present the alternative to the Single-Bottleneck Model.

Alternative Models

The class of alternative models that we consider here is that of models assuming that resources are divided between tasks (e.g., Kahneman, 1973; Kantowitz & Knight, 1976a; McLeod, 1977; Navon & Gopher, 1980; Norman & Bobrow, 1975; Wickens, 1980). For the sake of simplicity we assume that there is just one pool of resources. Let this be termed the *Single-Resource Model*.

Resources could be evenly allocated between tasks, regardless of the onset precedence of the tasks. On the other hand, some priority could be given to the first task. As long as that priority does not grant absolute exclusivity to the first task, it entails resource allocation rather than queuing, assuming some priority for the first task is not at all arbitrary. One, a defining property of divisible resources is that they are flexibly allocatable (Navon, 1985). Resource theorists, starting with Norman and Bobrow (1975), typically posited that. Thus, that resources can be evenly divided does not mean that they must be so. Two, there is sense in favoring the first task, if the instructions call for a particularly speeded response to it or if some tacit demand characteristics induce the subject to do that. Resource models would naturally allow the allocation of resources to be sensitive to such influences.

Assigning absolute priority to the first task would result in queuing. Queuing by strategy is qualitatively different from queuing by fiat or by necessity. Yet, since there is no easy way to empirically distinguish between the two, we focus here on pitting both against a model that assumes a positive, non-negligible allocation to the second task. Whenever we refer to the Single-Resource Model henceforth, this assumption is implied.

The issue to be examined here is to what extent RT data can help to decide between the Single-Bottleneck Model and the Single-Resource Model. Since the Single-Bottleneck Model may be actually regarded as the pole of a continuum spanned by all possible resource allocations within the Single-Resource Model, it is obvious that for any given imperfect degree of experimental power, the Single-Resource Model assuming any allocation ratio within some range starting at 1:0 could equally well account for data which are compatible with the Single-Bottleneck Model. Thus, the issue is evidently not whether the hypothesis of maximally uneven allocation (equivalent to strict queuing) can be accepted, but rather how broad the range of equally plausible allocation ratios is, hence which allocation ratios are ex-

cluded by the data. We show below that often the range is considerably broad, predominantly including the 9:1 ratio and sometimes even the 3:2 ratio, and furthermore that several types of data often cited in support of the Single-Bottleneck Model are compatible with *any* ratio.

The arguments we present are couched within a mathematical analysis. In Appendices A and B we present rigorously defined versions of the Single-Bottleneck Model and the Single-Resource Model, respectively, and derive predictions from both. We then compare the predictions of both models against data of the sort generally seen to support the Single-Bottleneck Model in order to see whether or to what extent the data selectively support it.⁴

⁴ Actually there are three more classes of alternative models. The first class bears a family resemblance to the Single-Bottleneck Model. Rather than positing a single serial processor, however, it posits the existence of a pool of serial processors. Let them be called servers, since the word "bottleneck" does not carry a very clear meaning in plural. Processes—both externally assigned and internally generated—queue for service by any free member of that pool. Internal processes are generated in an experimental session as well, but they are pushed down the queue stack following the onset of imperative stimuli for the assigned tasks. This model seems to be worthy of consideration because it appears to be the smallest modification of the Single-Bottleneck Model that is qualitative: Processing is not restricted to a single serial processor, but tasks do queue for service. Nonetheless, on the surface concurrence processing will be observed. This model seems to fall into category 3 in the list above (section "What Actually is the Main Issue?"). Let it be termed MSSM—short for Multiple Serial Servers Model. A second class subsumes models of outcome conflict (Navon, 1985; Navon & Miller, 1987). Let it be termed OCM—short for Outcome Conflict Models. The notion of outcome conflict assumes that processing is basically parallel, but claims the existence of internal variables called alterants that are modulated by processing, presumably in a task-specific manner, and in turn affect the rate of processing of concurrent tasks, again presumably in a task-specific manner. When the anticipated outcome conflict is intolerable, queuing may be used to avert it. This notion thus seems to fall, just like the Single-Resource Model does, in categories 3 or 6 in the list above (section "What Actually is the Main Issue?"). Outcome conflict is sometimes equated with content confusability or task similarity (e.g., Pashler, 1994a), perhaps because some early studies (Hirst & Kalmar, 1987; Navon & Miller, 1987) elected to demonstrate outcome conflict by the effect of confusability. However, the definition "output or side effect occasioned by the interfering event" (Navon, 1985, p. 137) is clearly broad enough to allow for conflict between nonsimilar events or processes. For example, a task whose processing overarouses the system (Navon, 1984) need not be similar to another task that suffers the consequences of that overarousal. Neither does outcome conflict necessarily entail switching between tasks to avoid it, as suggested by Tsang, Shaner, and Vidulich (1995). Subjects have other options—either to suffer the consequences or to lower the rate of concurrent processing to diminish cross-talk (Kinsbourne, 1981). In our view, outcome conflict models describe any interference that is not due to competition for some provision for processing. The third class subsumes models which assume strategic queuing of some operations to avoid anticipated interference. Let it be termed SPM—short for Strategic Postponement Models. An example is the EPIC model (Meyer & Kieras, 1997a, 1997b), which attributes response delays to suspension of motor processing to avoid motor conflict in the case that both responses resort to the same response modality. Meyer and Kieras actually state that each motor processor is serial or that its capacity can support only one movement at a time. However, their predictions would stay essentially the same if they assumed that motor processing could be parallel at a severe cost on precision, if not accuracy. In theory, central queuing could also be due to

PREDICTIONS FROM THE SINGLE-BOTTLENECK MODEL: ARE THEY SPECIFIC TO IT?

In the following seven sections we discuss seven major clusters of predictions from the Single-Bottleneck Model and examine to what extent they selectively support it. To make the discussion as intelligible as possible, we refer to the mathematical appendices only when necessary. We resort instead to verbal arguments, light or heavy as the case may be. We cite a number of relevant studies in each section and also compare the fit of the models with several of them. It should be borne in mind, however, that since our analysis is methodological, it focuses on diagnosticity, not on judging the present state of evidence.

Prediction 1: RT2 Depends Heavily on SOA

According to the Single-Bottleneck Model, the longer SOA is, the less likely is processing task 2 in the bottleneck to have to await the completion of processing task 1 in it. That is true regardless of wherever the bottleneck is and however short its processing time is.

Pashler (1993, 1994a, 1998) illustrates this rationale by the persuasive analogy of a single teller in the bank serving a queue of customers. If the teller managed to overcome some human tendencies (like adjusting his handling of a customer to the length of the line behind her), then the time a customer spends in the bank would be linearly related, with a slope of -1 , to the time that customer (later referred to as customer 2) lags after the previous customer (later referred to as customer 1) in entering the bank.

The -1 slope at the left end of the RT2–SOA curve is an obvious prediction from the Single-Bottleneck Model⁵ [see also Appendix A, Eq. (4)] because when the tasks must queue for processing by the bottleneck, the time the second task waits is obviously curtailed by one millisecond for every millisecond increase in the time S2 onset lags after S1 onset.

Actually, slopes close to -1 have not always been found (cf. review in Kahneman, 1973), but they are prevalent enough (cf. review in Pashler,

an attempt to avoid central conflict. Naturally, most predictions of models of this sort would be identical with the predictions of models of structural queuing, though they might differ with respect to locus. RT data are, unfortunately, indifferent to what drives the queuing. This class of models thus seems to fall in category 5 in the list above (see section “What Actually Is the Main Issue?”). It is beyond the scope of the present article to consider thoroughly models of the latter three classes. We do this with respect to MSSM and OCM in another article (Navon & Miller, 1998) and we just allude here to that analysis briefly under Discussion. As for SPM, it does not seem to sufficiently differ from the Single-Bottleneck Model to justify further analysis beyond that already done by Meyer and Kieras (1997a, 1997b) and by Ruthruff, Pashler, and Klaassen (1996). We comment on it briefly under Discussion.

⁵ Subject to one condition: That the encoding of S2 is not overly lengthy (see Appendix A). In most experimental studies that condition must have held.

1994a). Such slopes appear to constitute a very compelling piece of evidence in favor of the Single-Bottleneck Model.

So far, so good. However, is a large negative slope, even one that approaches -1 , an exclusive prediction from single-bottleneck models? Is it really incompatible with the Single-Resource Model?

An alternative account. Imagine a completely different sort of bank in which customers are serviced in parallel by a limited-capacity system.⁶ Precedence of arrival may or may not be given some priority, but if it is given, *any* allocation ratio may be chosen. This modified system is of course the analog of the Single-Resource Model. What is the effect of interarrival interval in such a system.

Suppose there are only two customers at the bank. When customer 2 overlaps with customer 1, the rate of her service consumption is slower. As soon as customer 1 finishes, that rate is at least doubled. That fortunate event is likely to occur sooner after the arrival of customer 2, the longer the latter lags after the arrival of customer 1. Hence, the longer the interarrival interval, the shorter the average time to complete servicing the second customer. By the same token, the negative effect of SOA on RT2 in the overlapping tasks paradigm is predicted to obtain in such a system as well, hence does not necessarily indicate either a refractory period or the existence of a single serial bottleneck.

Thus, the dependence of RT2 on SOA may also be accommodated by the Single-Resource Model. It might appear though, that the Single-Resource Model predicts a smaller effect of SOA than the Single-Bottleneck Model does. Is it really so? If so, how much smaller is the effect? Does the Single-Resource Model accommodate a -1 slope at the left end? How fast does the slope change?

The mathematical answer is surprisingly clear-cut: Regardless of what the allocation ratio might be, the slope is predicted to be the same as the Single-Bottleneck Model predicts it to be, namely -1 at the left end and progressively closer to 0 as SOA is increased [see Appendix B, Eq. (6)].

If this seems counter-intuitive, another way of putting it (following McLeod, 1977, and a personal communication by John Duncan) might help: There is some definite amount of work left to be done on both tasks at the moment that processing of S2 (at the limited stage) is ready to start. The

⁶ That is somewhat difficult to arrange if the services are provided by human tellers and there are fewer tellers than customers. However, suppose every customer is provided with a terminal, and services are provided by a powerful time sharing computer with time lags that are well below the human jnd. Processing time in that computer is thus a limited resource to be divided with no priorities between all customers presently at the premises. In passing, note that even in a more customary bank setting in which services are provided by a number of tellers, the larger the interval between the arrival of a customer and the arrival of the subsequent one, the less the latter would wait on the average for the first teller to be free (see Navon & Miller, 1998).

Single-Resource Model assumes that that work is done by a system having a total rate that is constant throughout a trial. Resource allocation affects the relative length of the overlap period (see Fig. 7 in Appendix B), but since both the total work yet to be done and the total rate are independent of the allocation ratio, their quotient, namely finishing time, must also be so. To the extent that R2 comes after R1, RT2 must thus be unaffected by the allocation ratio. That clearly applies also to the Single-Resource Model under the most extreme allocation ratio, namely 1:0. Since that is empirically indistinguishable from the Single-Bottleneck Model, it follows that both models make exactly the same prediction.

Prediction 2: RT1 Does Not Depend on SOA or Second Task Parameters

The Single-Bottleneck Model assumes that the processing of the second task is postponed, so that the first task is processed independently of it. Hence, RT1 should not depend on any characteristics of the second task, including the moment at which S2 is presented [see Appendix A, Eq. (1)].

Actually, the data do not strongly support this prediction. True, the effect of SOA on RT1 is generally negligible when the instructions favor the first task, although some significant effects have been reported (e.g., Smith, 1969). When subjects are asked to make both responses as quickly as possible, however, the effect of SOA on RT1 can be greater (e.g., a significant slope of .3 at the left of the curve in Pashler, 1991, Exp. 3). When the order of stimuli is blocked but subjects are not informed of the order, the slope can be even greater (about .5 in Pashler, 1990, Exp. 3).

RT1 was also reported to be affected to some degree by factors affecting the difficulty of the second task (e.g., Pashler, 1984; Smith, 1969).

These embarrassing effects on RT1 were explained (e.g., by Pashler, 1984; Pashler & Johnston, 1989) as following from a tendency of subjects to group their responses. If subjects do indeed often group their responses, that would explain the failure of some of these predictions to bear out: RT1 would positively depend on factors affecting the difficulty of the second task [for a formal analysis see Appendix D, Eq. (5)]. On the other hand, even when this assumption is used to augment the Single-Bottleneck Model, RT1 is still not expected to decrease with SOA but rather to slightly increase with it, unless a further, specific assumption is made about the frequency of grouping. Furthermore, it is doubtful that subjects use response grouping in a large enough percentage of trials to account quantitatively for observed effects on RTs (see Appendix D).

Similarly, the considerable increase in dual-task RT1 relative to single-task RT1 (e.g., Pashler, 1984; Smith, 1969) is not predicted by the Single-Bottleneck Model. The Single-Bottleneck Model proponents typically ascribed the discrepancy to response grouping or to the impact of the dual-task condition on the preparation for task 1, although they could not be sure that those could fully account for the effect.

The picture is further complicated by recent demonstrations of considerable sensitivity of RT1 to S1–R2 compatibility as well as R1–R2 compatibility observed under certain circumstances (Hommel, 1998; Logan & Schulkind, 2000). Those findings could be explained by positing parallel automatic processes of response retrieval that operate side by side with serial controlled processes or invite a forced serial strategy to obviate or overcome the response conflict they generate (cf. van der Heijden, 1981), still they are difficult to explain by the simple notion that selecting R2 does not start until the selection of R1 is completed, a notion that underlies the Single-Bottleneck Model (see, e.g., Fagot & Pashler, 1992; Pashler, 1994a). In addition, Logan & Schulkind found that the need to switch S–R mapping from task 1 to task 2 had similar effects on RT1 and RT2, which also cannot be easily accommodated by the Single-Bottleneck Model.

Thus, the Single-Bottleneck Model's clear-cut prediction of invariance of RT1 to SOA, task 2 factors, and even to the dual-task requirement itself has not been borne out. Explanations of RT1 effects in terms of response grouping or preparation can save the Single-Bottleneck Model from being rejected, but the fact remains that the Single-Bottleneck Model requires that imported help.

One aspect of the RT1 data are, however, quite in accord with the predictions of the Single-Bottleneck Model. Specifically, effects of these factors on RT1 are generally much smaller than effects on RT2. In contrast, the demands of the premotor stages of the first task (or the task that is processed faster; Ruthruff, Pashler, & Klaassen, 1996) affect RT2 as much as they affect RT1. The asymmetry is even more pronounced in the effects of SOA. This asymmetry appears to selectively support the Single-Bottleneck Model. Does it really?

An alternative account. Let us return to the bank analogy. The analog of the Single-Resource Model posits that the rate of service is inversely related to the number of customers to be serviced. Due to variability in service demand and in arrival time, temporal overlap may be partial; in other words, there may be periods during which a customer might enjoy exclusivity of service. When there are only two customers, those periods may be quite long.

Clearly, within this model interference goes in both ways, but the interference is not at all symmetric. The main reason is that the service demanded by customer 1 determines how long customer 2 would have to suffer reduced rate of service, whereas the service demanded by customer 2 determines only whether and how long she would still have to be serviced after the overlap period. Note also that when the service demanded by customer 2 is particularly modest, she might be serviced only during the overlap period, namely with a reduced rate throughout. That by definition cannot be true of customer 1. Thus, customer 1 often suffers less interference from customer 2.

Furthermore, the longer the interarrival interval is, the shorter the overlap period is. Hence, larger interarrival intervals will decrease service time for

both customers. However, since interference is asymmetric, that effect need not be the same on both customers. It would be the same if rate of servicing is reduced to the same extent for both customers. It would not, if some priority is given to precedence of arrival.

This conditional prediction is demonstrated in Appendix B ('Dual-task decrement'): The Single-Resource Model predicts that RT1 would be affected by the dual-task requirement, though often not to the same degree as RT2 is. It predicts that RT1 would hardly be affected by the demand of the second task for resource-limited processing, whereas RT2 would be strongly affected by the corresponding demand of the first task. It may also accommodate a smaller effect of SOA on RT1 than on RT2. However, it can accommodate a null effect of SOA on RT1 only at the cost of making either of two ancillary assumptions: that the demand of the first task is particularly small or that the proportion of resources allotted to it is particularly high.

Note, however, that the instance of the Single-Resource Model usually considered for that case makes some simplistic assumptions. One is that any task can use, when processed alone, all the resources available for processing both tasks concurrently. Another one is that all the resources are reallocated at the moment that task 2 is ready to use some. It could be, however, that task 1 uses from the outset just part of the pool and that task 2 can then use the remainder. Even when the residual amount of resources is not sufficient for the optimal processing of task 2, task 1 may be allowed to finish without reallocation for reasons such as sluggishness of reallocation (cf. footnote 2), confusability of outputs, and so on. In any event, the consequence of such partial queuing would be the same as that of strict queuing: At short SOA values, RT2 would heavily depend on SOA, but RT1 would not. Finally, the heavy load at short SOA values may drive subjects to expend a greater part of their potential capacity than they would at long SOA values, which might offset a small effect of release from load on RT1 if such existed.

All in all, the Single-Resource Model can accommodate effects indicative of task asymmetry. It accounts for effects on RT1, to the extent that those exist, better than the Single-Bottleneck Model does. Its simplistic version fails to account for a null effect of SOA, but a more realistic one would. Any version accommodates task compatibility effects on RT1 better than the Single-Bottleneck Model (and probably more economically than hybrid models like the one proposed by Hommel, 1998), since the Single-Resource Model assumes parallel processing. As for the version of the Single-Bottleneck Model augmented by grouping and task preparation, it fits the data that it fits by virtue of a subsidiary assumption rather than inherent ones.

Thus, unlike in the case of the prediction for the effect of SOA on RT2, here neither verbal reasoning nor mathematical analysis yields a simple answer. To compare the models quantitatively, on this prediction and others, we tested how the models fit actual data by running some computer simulations.

We elected to fit the results of Experiment 1 in Pashler and Johnston

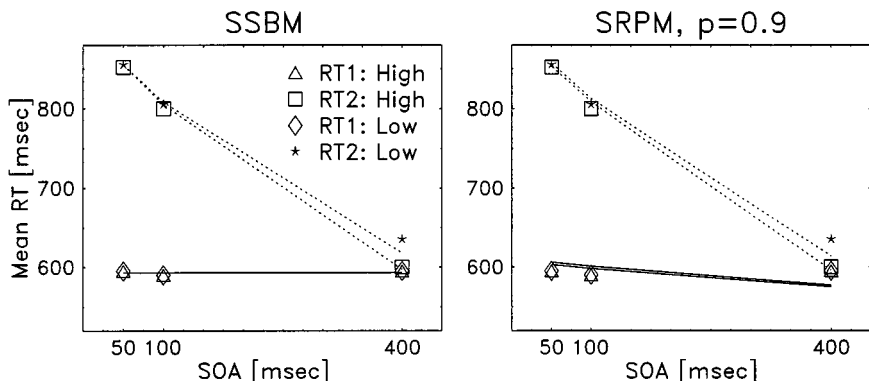


FIG. 1. The data of Pashler and Johnston (1989, Fig. 3) and the best fits of both models. The observed values are indicated by the points, and the model predictions are indicated by the lines. The figure presents mean RTs for correct responses as a function of task, SOA, and S2 intensity. "High" and "Low" refer to the levels of intensity of S2. SSBM and SRPM denote the Single-Bottleneck Model and the Single-Resource Model, respectively.

(1989) and of Experiment 1 in Smith (1969), each bearing on three of the phenomena discussed above and below, as well as the results of Experiment 1 in McCann and Johnston (1992) and Experiment 3 in Pashler (1991), each bearing on two of these phenomena. In view of the bias in favor of task 1 in the instructions, we have let the Single-Resource Model (in its simplistic version) assume $p = .9$, namely allocation of 90% of the resources to task 1. The details of the simulations, as well as brief descriptions of each of those experiments, are presented in Appendix C.

Predicted results are presented in Figs. 1–5 along with the actual results

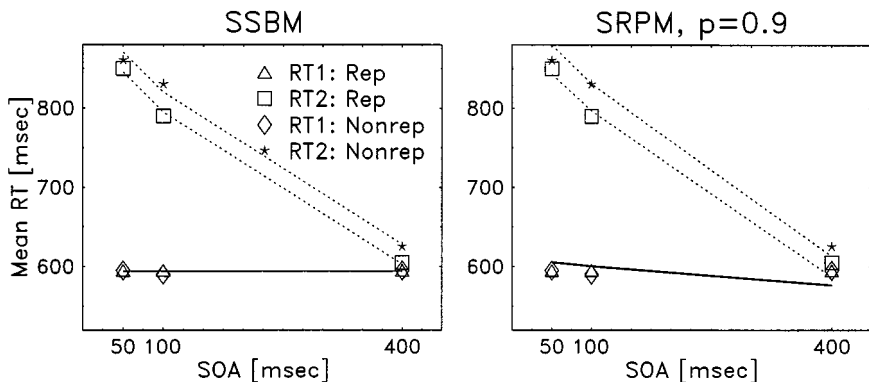


FIG. 2. The data of Pashler and Johnston (1989, Fig. 4) and the best fits of both models. The observed values are indicated by the points, and the model predictions are indicated by the lines. The figure presents mean RTs for correct responses as a function of task, SOA, and S2 repetition. "Rep" and "Nonrep" refer to the levels of repetition of S2.

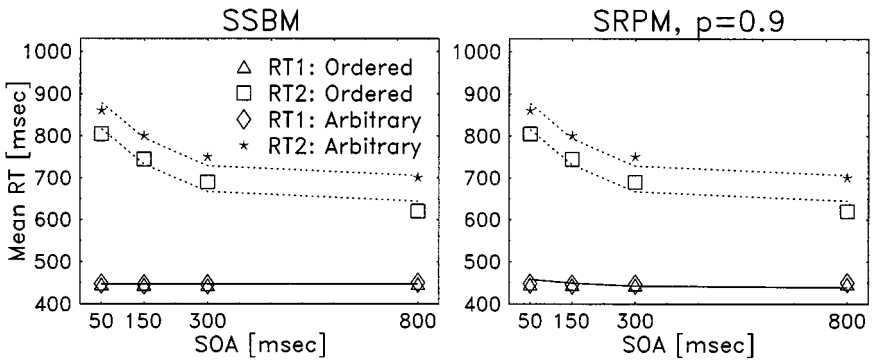


FIG. 3. The data of McCann and Johnston (1992, Fig. 3) and the best fits of both models. The observed values are indicated by the points, and the model predictions are indicated by the lines. The figure presents mean RTs for correct responses as a function of task, SOA, and difficulty of S2-R2 mapping. “Ordered” and “Arbitrary” refer to the possible mappings of S2 onto R2.

being fit. Goodness-of-fit values (RMS errors) are presented in Table 1A. As can be seen, as a whole the Single-Bottleneck Model does not fit the data clearly better than the Single-Resource Model with $p = .9$. The differences in favor of the Single-Bottleneck Model in RMS error (ranging from 4.5 to 8.2 with a median of 2.6) seem very slight compared with the RMS error measures for two obviously wrong models—one, the Single-Resource Model with $p = .1$, and two, a model that assumes unlimited parallel processing of both tasks (e.g., 102.0 and 71.9, respectively, for the fits to the data of Pashler & Johnston, 1989, shown in Fig. 3).

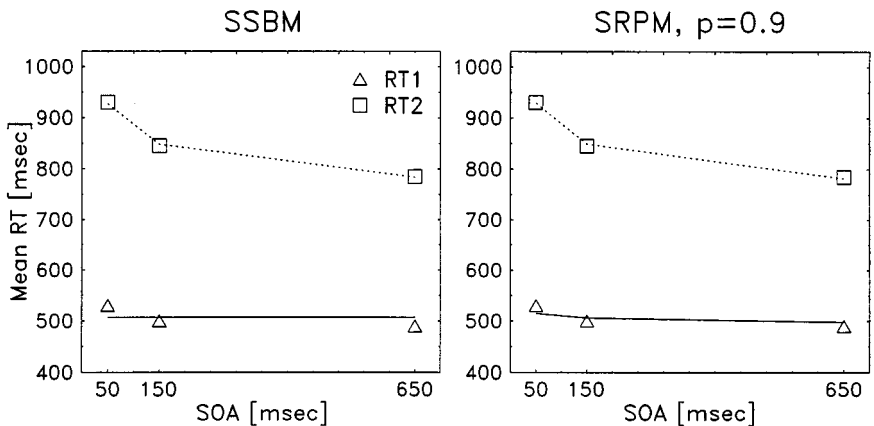


FIG. 4. The data of Pashler (1991, Fig. 8) and the best fits of both models. The observed values are indicated by the points, and the model predictions are indicated by the lines. The figure presents mean RTs for correct responses as a function of task and SOA.

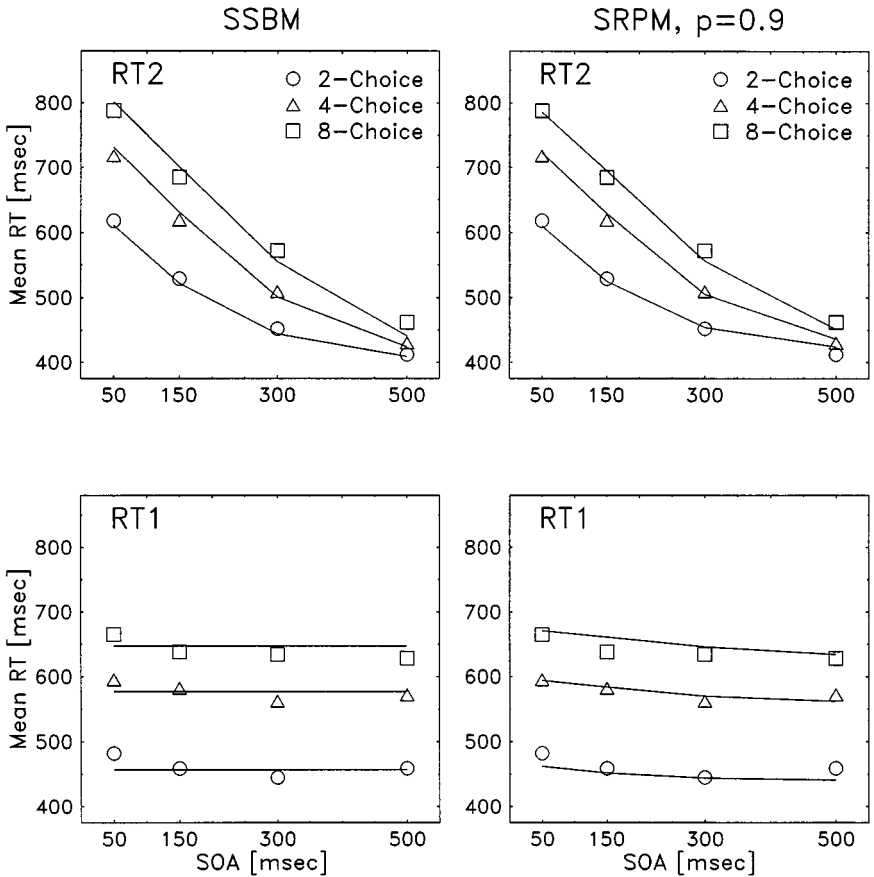


FIG. 5. The data of Smith (1969, Figs. 1 and 2) and the best fits of both models. The observed values are indicated by the points, and the model predictions are indicated by the lines. The figure presents mean RTs for correct responses as a function of task, SOA and task 1 difficulty. "2-choice," "4-choice," and "8-choice" refer to the number of alternative stimuli serving as S1.

Furthermore, in the single case in which the instructions were relatively neutral (Experiment 3 in Pashler, 1991), we have added simulations of the Single-Resource Model with different values of p . As can be seen in Table 1B, the fit of the Single-Bottleneck Model (RMS error = 12.1) was somewhat surpassed by fits of the Single-Resource Model with values of p ranging from about .63 to .95 (with an optimum around .8, where RMS error = 3.4). In comparison, the two obviously wrong models yielded, respectively, RMS errors of 75.1 and 43.7.

In the simulations we ran, we did not attempt to fit data concerning the effects of the dual-task requirement and the difficulty of the second task on

TABLE 1

(A) Parameter Estimates and Goodness-of-Fit Values (RMS Error) for All Models and Data Sets and (B) Goodness-of-Fit Values for the Single-Resource Model with Various Values of p to the Data of Pashler (1991, Fig. 8)

A. Parameter estimates and goodness-of-fit values

Data Set	RMSe	g1	g2	c1	c2
Single-Bottleneck Model					
McCann and Johnston (Fig. 3)	12.0	224	353	174	246
Pashler (1991, Fig. 8)	12.1	195	453	263	286
Pashler and Johnston (Fig. 3)	5.7	393	168	155	300
Pashler and Johnston (Fig. 4)	4.8	352	376	196	125
Smith (1969, Figs. 1 and 2)	13.0	259	55	150	300
Single-Resource Model, $p = .90$					
McCann and Johnston (Fig. 3)	13.2	223	303	168	295
Pashler (1991, Fig. 8)	7.6	198	461	252	277
Pashler and Johnston (Fig. 3)	11.8	457	127	61	300
Pashler and Johnston (Fig. 4)	13.0	429	136	91	300
Smith (1969, Figs. 1 and 2)	10.4	241	107	152	267

B. Goodness of fit for the Single-Resource Model

p	RMSe
0.50	22.1
0.55	17.8
0.60	13.7
0.65	9.8
0.70	6.3
0.75	3.6
0.80	3.4
0.85	5.3
0.90	7.6
0.95	10.5

RT1, since it is clear that these are incongruent with the Single-Bottleneck Model unless it is augmented by assumptions on response grouping and so on. We did fit the slight effect of SOA on RT1 (reported, e.g., by McCann & Johnston, 1992; Pashler & Johnston, 1989). It can be seen in Figs. 1–3 that the Single-Resource Model can fit this pattern by assuming resource allocations that are heavily biased in favor of the first task. On the other hand, the Single-Bottleneck Model fits Smith's findings slightly worse than the Single-Resource Model does (see Fig. 5) and is certainly embarrassed by results of studies in which the instructions do not so clearly favor the first task (Pashler, 1991, Exp. 3; see Fig. 4 above).

In sum, while extensively studied effects on RT2 data are barely pertinent for the issue, RT1 data are definitely more so. Still, interpreting them depends

on one's perspective. On the one hand, they clearly differ from RT2 data, which the Single-Bottleneck Model predicts. On the other hand, they are not as independent of SOA and other factors as the Single-Bottleneck Model predicts them to be. Even a small effect could be quite damaging to a model that predicts a null effect. One may rightfully claim that while the cup is not full, neither is it empty.

Prediction 3: RT1 and RT2 Are Correlated

In the overlapping tasks paradigm, RT1 and RT2 are typically found to be highly correlated, at least at low SOA values (e.g., Borger, 1963; Gottsdanker & Way, 1966; Pashler & Johnston, 1989). This is surely compatible with the Single-Bottleneck Model (see also Appendix A, "Correlation"): The longer it takes the bottleneck to process the first task, the longer the second task has to wait for its turn.

However, magnitudes of a correlation are notorious for being hard to predict. First, they must be affected by the tendency for response grouping (see direct evidence in Pashler & Johnston, 1989). Second, they could be affected by many nuisance variables. In this case, the most prominent nuisance variables are arousal, interest, and distraction (Navon, 1990a). Those probably fluctuate in the course of an experimental session and affect RT1 and RT2 in the same manner.

Hence, without a valid quantification of the influence of such variables, a high correlation coefficient in itself is barely informative. How it is affected by experimentally relevant parameters might be more illuminating. For example, covariation that is due to nuisance variables need not be very sensitive to SOA, whereas covariation due to queuing would decline with it (see also Appendix A, "Arousal effects"). So would covariation due to response grouping, if grouping were more frequent at smaller SOAs (see Appendix D). More specifically, if the correlation was due only to the effect of task 1 parameters on RT2, it would be very low at an SOA level in which task 2 would rarely have to wait for the bottleneck (see Appendix A, "How long is long?", for ways to determine that an SOA level satisfies this condition).

Studies that measured correlation coefficients⁷ often report them not only within the range of SOA at which mean RT2 declines, but also for SOA values that are far beyond that. But even within the relevant range, the pattern

⁷ In some other studies (e.g., Pashler, 1989; Pashler & O'Brien, 1993) a different association measure was used. They reported a large reduction with SOA in the slopes of curves plotting mean RT2 within RT1 quintiles. Those, however, are not measures of correlation but rather coefficients of regression by conditional means. They, like any other regression coefficient in raw scores, are affected not only by the correlation but by the variances as well. Now, there must be some decline in the variance of RT2 scores with SOA, since mean RT2 declines rapidly with SOA and means and variances of RTs tend to be correlated (see Luce, 1986). Hence, the reported reduction in slopes is probably inflated considerably. To illustrate, a decline of the slope from 1.0 to .4 might reflect a decline in Pearson r from .9 to .65.

emerging is somewhat equivocal. In view of the high variance in subjects' coefficients and the small number of subjects used in some old studies, that is hardly surprising. To illustrate, Davis (1959) reported an RT1–RT2 correlation at an SOA of 50 ms for only one subject (its value is .77), Gottsdanker and Way (1966) reported a reduction in correlation (averaged over 8 subjects) from .39 to .19 as SOA increased from 50 to 800 ms, and Karlin and Kestenbaum (1968) reported a reduction (averaged over 12 subjects from three experiments) from .69 to .33 as SOA increased from 90 to 590 ms. In more modern research, Pashler and Johnston (1989, Exp. 1) reported a modest reduction in the mean Pearson coefficient (averaged over 36 subjects)—from .78 to .64 as SOA increased from 50 to 400 ms.

An alternative account. The predictions of the Single-Resource Model with respect to the RT1–RT2 correlation are much less intuitive. Competition for resources appears to induce a negative correlation. However, that would be true only if allocation proportion was allowed to vary across the set of trials over which the correlation was computed (e.g., within a given SOA). If this proportion were instead fixed, the correlation would be determined by other factors. For example, nuisance factors such as arousal, interest, or distraction would tend to make it positive, as noted above. More important, the demand of the first task affects the length of the period during which the second task is processed at a reduced rate. Hence, the variability of the demand of the first task contributes to a positive correlation (see Appendix B, "Correlation"). Hence, RT1 and RT2 might plausibly be positively correlated, even though task 1 and task 2 compete for processing resources.

Within the Single-Resource Model, both RT1 and RT2 decrease as SOA increases. If SOA variability had been manipulated, then, it would have tended to produce a positive correlation. Finally, here too, when SOA is large enough that the tasks do not truly overlap, the correlation would be limited to that induced by the common nuisance factors.

Thus, both models can account for a positive correlation between RT1 and RT2. Both, particularly the Single-Bottleneck Model, predict it to decline considerably with SOA. The fact that the decline is only moderate is better accounted for by attributing the correlation mainly to common nuisance factors such as arousal than by attributing it to queuing. The bottom line seems to be that the RT1–RT2 correlation is not diagnostic.

*Prediction 4: The Demand of Selecting R2 Is Additive with SOA,
but the Effect of S2 Discriminability Is Washed Out*

Stimulus discriminability (or intensity) normally affects reaction time. However, in the overlapping tasks paradigm, the effect of the discriminability of the second-task stimulus (S2) is smaller than it would be in a single-task condition (de Jong, 1993; Pashler, 1984; Pashler & Johnston, 1989). Moreover, this discriminability effect is especially small when SOA is short.

Thus, the effect of S2 discriminability is said to “wash out” at short SOAs in the overlapping tasks paradigm.

The Single-Bottleneck Model explains this washout in a straightforward way (see, e.g., Pashler, 1994a, p. 224): Encoding S2 can often be done while the bottleneck is still busy selecting the response for S1. In that case, RT2 would be affected by the demand of encoding and responding to S1 but be insensitive to the demand of encoding S2 [see also Appendix A, Eq. (3b)]. On the other hand, the Single-Bottleneck Model predicts that the demand of selecting the response for task 2 would affect RT2 to the same extent irrespective of SOA because response selection is done in the bottleneck (cf. Pashler, 1994a, p. 224).

Actually, the very same rationale (sometimes called locus-of-slack approach) is used by some Single-Bottleneck Model proponents to argue that the bottleneck is central (e.g., Pashler, 1984). If it were localized instead in the response initiation stage, the demand of response selection would interact with the dual-task requirement or with SOA, just as the demand of encoding does. The fact that the effects of those factors are typically additive (see review in Pashler & Johnston, 1998; but see Hawkins, Rodriguez, & Reicher, 1979; Karlin & Kastenbaum, 1968; Shumacher et al., 1999) seems to suggest that the bottleneck has a central locus (McCann & Johnston, 1992; Pashler, 1994a; Pashler & Johnston, 1989).

An alternative account. The Single-Resource Model is clearly quite compatible with the additivity of SOA with the demand of selecting R2. It simply predicts it for cases in which there is some overlap in processing periods [Appendix B, Eq. (6), cond2]. On the other hand, the Single-Resource Model has often been assumed to be embarrassed by the washout of the S2 intensity effect (e.g., McCann & Johnston, 1992; Pashler, 1984, 1994a). The reason seems to be that proponents of the Single-Bottleneck Model relied on McLeod's (1977) argument that resource models predict an overadditive interaction between every difficulty manipulation and the dual-task requirement (or SOA). There are two faults in this logic, however. One, overadditive interactions are predicted only under certain conditions (see below discussion of prediction 5; also see Navon, 1984, for an extensive discussion), and it is not clear whether those conditions are met in the experiments that demonstrate the washout effect. Two, overadditive interactions are predicted when the affected processes compete for resources, but resources might not be required at all stages. In particular, encoding may be automatic (e.g., Keele, 1973; Posner, 1978) or data-limited (Norman & Bobrow, 1975) or require a kind of resource that is different from the one required for response selection (Johnston, McCann, & Remington, 1995; Navon & Gopher, 1979; Wickens, 1984). Unless that possibility is ruled out, the absence of an overadditive interaction should not be taken as evidence against the Single-Resource Model.

If for any reason encoding is not limited by the resources that are required

for response selection, a little rethinking might convince us that the Single-Resource Model makes the same prediction, namely that in the dual-task condition, when SOA is small, the intensity of S2 will not affect RT2 at all (cf. Pashler, 1994b). The reason is that during the stage of S2 encoding, processing the first task proceeds as usual. Hence, that stage serves, much as SOA does, to shorten the period during which the second task is processed at a reduced rate. Unlike SOA, though, that stage does contribute to RT2. It is shown in Appendix B [Eqs. (4b) and (6)] that, under plausible conditions, the positive and negative effects of increasing S2 encoding time cancel each other completely, leading to a washout of the S2 effect at short SOAs.

Thus, both models can account both for the additivity of SOA with the difficulty of selecting R2 and for the washout of the effect of S2 intensity on RT2 in a dual-task condition. Thus, the Single-Bottleneck Model is not selectively supported by this pattern of results. This conclusion is further strengthened by the adequacy of the fits in the simulation results presented in Figs. 1–2.

Prediction 5: Effect of the Difficulty of the First Task Is Not Modulated by the Dual-Task Requirement

The Single-Bottleneck Model predicts that the difficulty of the first task (to be accurate, of the bottleneck processing of the first task) would affect RT1 in the dual-task condition as much as it does in a single-task condition. This simply falls out of the strict precedence for processing the first task. The equation for RT1 under the Single-Bottleneck Model [see Appendix A, Eq. (1)] is identical with the one that would apply in a single-task situation.

As mentioned above, RT1 was actually found to be affected by the dual-task situation. Nonetheless, the dual-task requirement did not interact with the effect of the difficulty of the first task (Pashler, 1984).

An alternative account. There is no unique prediction from resource theory about the interaction of difficulty and dual-task requirement (see Navon, 1984). Nonetheless, the Single-Resource Model has been argued to be embarrassed by the additivity of task 1 difficulty with dual-task requirement (see, e.g., Pashler, 1994a).

Apparently, this is because there is a clear-cut prediction from a simple model of resources which assumes (a) performance of both tasks depends on resources from a single pool, (b) no data limit is encroached, (c) the difficulty manipulation affects load rather than rate of processing (e.g., McLeod, 1977; Navon & Gopher, 1980), and (d) the difficulty manipulation affects a stage that is resource limited or a stage that precedes the resource limited stage. Such a model predicts [see Appendix B, Eq. (5)] that the effect of the demand of the first task on RT1 (specifically, up to response initiation) would be inflated in the dual-task condition (except for trials in which SOA is large or the demand is particularly low). The reason is simply that any difference in the amounts of time to complete a large task versus a small

one is magnified by a reduction in the supply of resources to accomplish the work.

It is far from clear, however, that all of these assumptions have been met in any study yielding the problematic additivity. Consider, for example, the issue of the difficulty manipulation. It often takes some speculation to assert that it affects load (see Navon, 1984). Even a factor such as display size that appears to influence the amount of encoding might possibly affect the search process in a manner that has nothing to do with the number of required operations, as proposed by Duncan (1980). If it affected processing rate rather than processing load, resource theory no longer necessarily predicts an interaction.

Furthermore, the predicted interaction depends on the specific resource allocation in the dual-task condition. The effect is expected to be small, and probably hard to detect, when the first task is assigned high priority [see Appendix B, Eq. (5)]. As argued, it is quite likely that in the overlapping tasks paradigm, the first task gets most of the resources because of the inherent demand characteristics as well as the instructions. Increasing the difficulty of the first task might also lead to an increase in its share of resources, which would tend to counteract the effect in question.

Thus, the additivity of the dual-task requirement and the difficulty of the first task is predicted by the Single-Bottleneck Model, but it is not actually incompatible with the Single-Resource Model.

Prediction 6: Difficulty of the First Task Affects RT1 and RT2 Equally as Long as Stages at or before the Bottleneck Are Tapped

The Single-Bottleneck Model predicts that in a dual-task situation the difficulty of the stages of processing the first task up to the bottleneck stage would affect both RT1 and RT2 to the same extent (see, e.g., Pashler, 1994a, p. 224). Like prediction 5, this prediction also falls out of the strict precedence for processing the first task. From the equation for RT2 under the Single-Bottleneck Model [see Appendix A, Eq. (3b)], it can be seen that as long as bottleneck processing of task 2 is postponed to some extent (cond2), RT2 is affected by the durations of the first two stages of task 1 as much as RT1 is [see Appendix A, Eq. (1)], so that any systematic increase in the duration of the third stage would affect both latencies to the same extent.

That prediction has been generally borne out. As can be seen in Fig. 5, Smith (1969) obtained nearly equal effects of task 1 difficulty on RT1 and RT2. Pashler (1994a, p. 224) cites three more cases, and Pashler (personal communication) noted that there are at least seven more.

The Single-Bottleneck Model also predicts (see, e.g., Pashler, 1994a, p. 224) that the demand of postbottleneck stages of task 1 processing would affect only RT1 but not RT2. The reason is that according to the model, RT1 is affected by the duration of that stage [Appendix A, Eq. (1)] but RT2 is

not [Appendix A, Eq. (2)]. Testing this prediction requires the assumption that the bottleneck applies no later than the process of response selection. That is probably true only in part, since attempts to confirm the prediction were only partly successful (results of Pashler & Christian, reported in Pashler, 1994a).

Can these findings be accommodated by divisible resource models? As before, we will consider a simple version, namely the Single-Resource Model, and grant as well that the difficulty manipulation affects a stage that is resource limited.

An alternative account. It is easy to see [see Appendix B, Eqs. (1), (4a), and (4b)] that the Single-Resource Model directly predicts that the demand of the latest stage of the first task would affect only RT1. A closer look is needed to examine the prediction of the Single-Resource Model with respect to the effects of the difficulty of encoding S1 or of selecting a response to it.

First, let us consider a case in which the tasks are processed concurrently with no temporal advantage to either task. That could happen either when $SOA = 0$ and strategic queuing was not opted for, or when $SOA > 0$ but subjects opted to synchronize the onsets of processing in both tasks. In that case, the prediction of resource theory depends absolutely on the relative importance of the two tasks. Specifically, Navon and Gopher (1979, p. 220) claim that the task whose difficulty is manipulated is likely to be affected more than the concurrent task, except when its performance is protected as when it is defined by the experimenter as primary. It does not seem implausible to conjecture that task 1 is being protected in some sense in most experiments using the overlapping tasks paradigm. It would not be surprising in that case, if the difficulty of task 1 affected RT2 just as much as RT1.

Now let us examine the case in which $SOA > 0$ and subjects start bottleneck processing of S2 only after the processing of S1 has already started. In that case, it can be seen in Appendix B [Eqs. (5) and (6)] that as long as bottleneck processing of task 2 is postponed to some extent (cond2), the effect of the demand for resources by response selection processing of task 1 on RT1 is greater than its effect on RT2. Thus, a systematic increase in that demand would delay RT1 more than it would delay RT2.

However, how easy would it be to demonstrate such an interaction? The magnitude of the discrepancy of effects depends on p , the proportion of resources allocated to the first task. For example, when $p = .5$, the effect on RT1 is expected to be twice as large as the effect on RT2. When $p = .9$, on the other hand, the former is expected to be only 10% larger than the latter. As argued above, it is quite likely that the allocation is biased in favor of task 1, because of the inherent demand characteristics if not because of the instructions, and that increasing its difficulty might even lead to an increase in its share of resources.

It remains to examine whether the experimental data cannot be fit by the Single-Resource Model as well as they are fit by the Single-Bottleneck

Model. In the set of results we used, Smith (1969) was the only one who reported the effects of task 1 difficulty on both RT1 and RT2. As can be seen in Fig. 5 and Table 1, the fit of the Single-Resource Model with $p = .9$ to her data is about the same as (and even slightly better than) the fit of Single-Bottleneck Model.

Thus, a lack of interaction between task and the difficulty of task 1 is predicted by the Single-Bottleneck Model, but is quite comfortably accommodated by the Single-Resource Model, unless one insists on confining it in the straitjacket of even allocation.

Prediction 7: IRI Does Not Depend on SOA

The Single-Bottleneck Model predicts that at its lower range SOA would have no effect at all on the interval between the two responses (IRI) which equals $RT2 - RT1 + SOA$. The reason is that as long as the second task has to wait for the termination of bottleneck processing of the first task, the exact time it waits clearly does not matter [see also Appendix A, Eq. (5b)]. The exact magnitude of SOA would affect the time the second task has to wait for its turn, but since the moment at which its turn arrives depends only on the processing of the first task, the interval between the two responses would not be affected at all.

In fact, that prediction is not generally borne out. Kahneman (1973) was the first to show a clear positive slope, contrary to the prediction of the Single-Bottleneck Model, at the left of the IRI-SOA curve by reanalyzing Smith's (1969) data. Such a slope seems to exist to various extents in quite a few sets of data (see, e.g., McCann & Johnston, 1992, Fig. 1), though not in all (see, e.g., Pashler & Johnston, 1989, Fig. 2).

Interestingly, the Single-Bottleneck Model cannot be rescued here by just assuming response grouping. Grouping predicts, of course, very low IRIs, but more importantly, a flat function of IRI versus SOA [see Appendix D, Eq. (7)]. It must be further assumed that subjects tend to group more when SOA is short (Welford, 1967). Since IRI is shorter when grouping takes place, the effect follows [cf. Appendix D, Eq. (8) and the following discussion]. However, this doubly augmented version of the Single-Bottleneck Model actually predicts a continuously accelerated slope and probably a flat region at the very left end of the curve [cf. Appendix D, Eq. (6), and following discussion]. That does not appear to be the case in Kahneman's (1973) plot of Smith's (1969) data. Unfortunately, proponents of the Single-Bottleneck Model seldom report this aspect of their data.

It is in fact quite important that the data do not support this prediction because it is not shared by the Single-Resource Model.

An alternative account. The Single-Resource Model plainly predicts a positive slope. The larger SOA is, the shorter the overlap period during which both tasks use resources (OVP in Fig. 7). When OVP is shorter, less of the demanded processing of task 2 is done during it and more is left to be done

during the subsequent period (OT2P in Fig. 7). Thus, SOA directly increases the lag of R2 after R1 [see also Appendix B, Eqs. (7a)–(7c)].

It should be noted, however, that this prediction is redundant with the conjunction of two other ones. By the definition of IRI, its dependence on SOA is positively related both to the deviation of the slope of the RT2–SOA curve from -1 and to the deviation of the slope of the RT1–SOA curve from 0 . Since the former is predicted by both models, it follows that the diagnostic value of the IRI–SOA curve for discriminating between the Single-Bottleneck Model and the Single-Resource Model is equivalent to the diagnostic value of the RT1–SOA curve. On the other hand, a decrease of mean RT1 with SOA appears unimpressive when compared to the dramatic decline of RT2 with SOA. Such masking by context is less likely to color the conclusion made from inspecting the slope of the IRI–SOA.

Thus, the Single-Bottleneck Model is embarrassed by the finding of a positive IRI–SOA slope which it clearly does not predict, whereas the Single-Resource Model easily accommodates it.

A Brief Overview of Predictions

The predictions of the two models with respect to 15 effects discussed in the past seven sections are summarized in Table 2.

TABLE 2
Summary of the Predictions of the Two Models

Prediction for the effect of	Model	
	Single-Bottleneck	Single-Resource
DTR on RT2	Large	Large
T1 difficulty on RT2	Large	Large
Slope at left of RT2–SOA curve	-1	-1
DTR on RT1	Null	Positive, but small when p is large
T2 difficulty on RT1	Null	Null
SOA on RT1	Null	Positive, but small when p is large
T1–T2 compatibility on RT1	Null	Positive
RT1–RT2 correlation at short SOA	Positive	Positive
RT1–RT2 correlation at long SOA	Zero, unless	common nuisance factors are assumed
S2 discriminability on RT2 at short SOA	Washout	Washout
SOA and demand of selecting R2 on RT2	Additive	Additive
DTR on effect of T1 difficulty on RT1	Null	Probably inflated
Difference between effects of T1 difficulty on RT1 and RT2	Null	Positive, but small when p is large
Difficulty of latest stage of T1 on RT2	Null	Null
Slope at left of IRI–SOA curve	0	Positive

Note. DTR is used to abbreviate dual-task requirement, T1 to abbreviate task 1, and T2 to abbreviate task 2.

The table, as well as the above reasoning, makes it clear that there is much less qualitative difference between predictions made by the two models than might be expected. It seems that on the whole both types of task interference behave more or less alike with respect to these predictions, surprising as that might seem. Some of the predictions are quantitatively identical for the two models regardless of the allocation ratio assumed by the Single-Resource Model. Most prominent of these is the prediction about the slope of the RT2–SOA curve (see section on prediction 1), often used to illustrate how effective a single bottleneck can be. Four other types of prediction (see sections on predictions 2, 5, 6, and 7) differ considerably only when the Single-Resource Model assumes allocation ratios that are about even, which is dubious given the emphasis on task 1 in the paradigm.

In the computer simulations we tried to fit only data that are relatively more problematic for the Single-Resource Model. The fits (see Figs. 1–5 and Table 1) demonstrate that on the whole the quantitative difference is very small, provided that the Single-Resource Model is set to allocation ratios that are quite reasonable for the circumstances. Actually, for data observed in an experiment where the instructions were neutral, the fit of the Single-Resource Model with allocation ratios larger than 3:2 was even somewhat better (see Table 1B).

One might contend that the comparison is unfair, because the Single-Resource Model enjoys the advantage of having a flexible allocation ratio parameter. That kind of argument would have been quite sound if either of the models had been concocted post hoc. However, both models had been originally formulated as well-reasoned hypotheses about the cause of task interference much before all the data in question have been collected.

Furthermore, the ‘extra’ parameter of the Single-Resource Model is not simply a luxury that the Single-Bottleneck Model spares. The Single-Bottleneck Model positively denies the existence of that parameter by assuming that concurrence is not feasible. It is not clear that making an assumption to spare a parameter is more parsimonious than adding a parameter to spare an assumption.

The Single-Resource Model is thus entitled to compete with the Single-Bottleneck Model in accounting for any given result, provided that it makes prior commitment about the plausible range of allocation ratios in the experimental situation. Thus, the crucial question is this: In view of the demand characteristics of the overlapping tasks paradigm, is it unreasonable for the Single-Resource Model to assume an allocation ratio of 3:2 (see Table 1B)? And when there is in addition an explicit bias in instructions, actually an invitation for serial processing, is it unreasonable to assume an allocation ratio of 9:1 (see Table 1A)?

Proponents of the Single-Bottleneck Model might tend to give a positive answer (see, e.g., Pashler, 1994a, p. 220), presumably because they are used to attributing to resource theory the assumption that resources are allocated evenly or by task demand. While it may not be wrong to attribute that as-

sumption to early notions of capacity sharing that typically kept mute about the subject (e.g., Broadbent, 1971; Kahneman, 1973; Keele, 1973; Kerr, 1973; Moray, 1967; Posner & Boies, 1971), forcing it on modern resource theory (e.g., Gopher, 1994; Kantowitz & Knight, 1976a; Kinchla, 1980; Navon, 1984; Navon & Gopher, 1979, 1980; Norman & Bobrow, 1975; Wickens, 1980, 1984) is plainly inaccurate.

Since it is not unreasonable to assume uneven allocation ratios in the overlapping tasks paradigm, the good fit to the data achieved by an instance of the Single-Resource Model assuming high priority for task 1 casts doubt on the hypothesis that they result from strict queuing. One might contend that an allocation ratio of 9:1 is close enough to queuing. Alas, the issue we set out from, as might be recalled, has not been how much emphasis is put on task 1 but rather whether concurrent processing is feasible. Since the Single-Bottleneck Model does not allow concurrence at all, even data that suggest allocation of 10% of the resources to task 2 would embarrass it, and data that do not empirically distinguish between this possibility and queuing would not selectively support it.

Finally, the assumption of uneven allocation was required to account for only three types of findings that appeared to support the Single-Bottleneck Model (specifically, the relative insensitivity of RT1 to SOA, the additivity of task with the difficulty of task 1, and the additivity of task 1 difficulty with the dual-task requirement). The remaining findings are compatible with the Single-Resource Model *under all allocation ratios*.

In sum, a number of effects supposedly supporting the Single-Bottleneck Model do not selectively do so. Is there evidence, obtained in the overlapping tasks paradigm, that even embarrass it? Notwithstanding the similarity between the predictions of the models, some phenomena—ironically, not the ones most frequently studied until recently—are more diagnostic than others. The Single-Bottleneck Model cannot easily accommodate effects of the existence of a second task, much less so of its complexity or timing, on RT1. The Single-Bottleneck Model is equally embarrassed by a positive slope at the left of the IRI–SOA curve.

Proponents of the Single-Bottleneck Model have tried to redress these unpalatable consequences mainly by invoking two auxiliary assumptions—that responses are often grouped and that a task may affect the preparation needed for performing the other task. Both assumptions are plausible, but augmenting the Single-Bottleneck Model by them poses some severe problems of testability. We elaborate on those problems in Appendices D and E.

DISCUSSION

The analysis presented here demonstrates that most of the typically reported types of result from the overlapping tasks paradigm do not selectively support the Single-Bottleneck Model over the Single-Resource Model.

Our analysis was carried out in two levels, model-based reasoning and

precise predictions calculated by computer simulations. The fits we obtained by the latter method add to the ones already made by Meyer and Kieras (1997b) using the EPIC architecture, yet another alternative to the Single-Bottleneck Model. Clearly then, the effects considered here are compatible with a variety of models rather than being evidence uniquely in support of the Single-Bottleneck Model.

Note that we have taken a different approach than Meyer and Kieras did. EPIC is actually a queuing model. Although it does not assume a structural bottleneck, it deliberately postpones some processing to avoid conflict or to secure desired response order. Whichever is the cause, queuing is the consequence. Being a queuing model, EPIC is bound to predict marked asymmetry of tasks and high sensitivity to SOA. In contrast, the Single-Resource Model does not assume queuing by fiat⁸ but does assume some real constraint on concurrent processing. Thus, the indeterminacy we found adds to the one pointed to by Meyer and Kieras. In another article (Navon & Miller, 1998), we extended the analysis presented here to two other models of task interference—a specific outcome conflict model and a specific queuing model that assumes two servers rather than one. The fits of these models with the same data were in general as good as the fits of the Single-Bottleneck Model and the Single-Resource Model.

How can one explain the fact that the alternative models make similar predictions? There seem to be some constraints built into the situation tapped by the paradigm, considering that it results in some interference: (a) Interference of any sort must grow rapidly with greater temporal overlap. (b) Interference is not symmetric. Asynchrony of the onset of processing is likely to give an advantage to the earlier task and (c) so would its earlier finishing time. (d) There are many possible reasons for cross-trial correlation between performance scores in the two tasks. (e) A particular locus of interference (e.g., central) can easily be accommodated by any model of interference.

The rest depends totally on the parameters. Our simulations indicate that with the proper adjustment of parameters of the distributions of stage components, each of the models can fit the data quite well. Thus, apart from ruling out limit-free parallel processing, the data do not seem to seriously embarrass known alternative models.

⁸ The Single-Resource Model is, actually, quite consonant with strategic queuing in the overlapping tasks paradigm, especially where the instructions stress speed of responding to task 1. Since RT2 is unaffected by the allocation ratio [see Appendix B, Eq. (6), or the argument presented under “Prediction 1: RT2 Depends Heavily on SOA”], subjects could shorten RT1 without any cost to RT2 by opting for strategic queuing. Adding to that the cost of cross-talk, queuing seems to be preferable (cf. Logan & Gordon, 2001). On the other hand, to the extent that performance depends on subjects’ deliberate planning, it is doubtful that they are always sophisticated enough to realize all the facts that bear on the optimization problem and apply the consequence to their behavior (if induction from other domains, probability matching for one, is valid here).

That indeterminacy is due mostly to the experimental situation. Theories that are clearly discrepant sometimes make similar predictions for a broad range of situations, so that a very specific one is required for testing between them.⁹

While for most practical purposes, queuing provides a satisfactory working model of behavior in the overlapping tasks paradigm, the indeterminacy shown here casts doubt on whether that queuing is strict. While that possibility is not a priori implausible, the data add little to substantiate it. But even supposing it was substantiated, that would not help to confirm the theoretical implication claimed by proponents of the Single-Bottleneck Model, namely the existence of a single bottleneck causing most dual-task interference, in view of the bias in the paradigm.

How Relevant Is the Paradigm Anyway?

Despite the doubts cast here on the interpretation of findings in the overlapping tasks paradigm, one might argue that they support strict queuing at least as well as resource-limited concurrent processing and expect that with some more experimental progress—perhaps even better. After all, strict queuing is a quite reasonable strategy for the paradigm. However, the issue we set out to examine (see section “What Actually Is the Main Issue?” above) was more general than determining whether the single-bottleneck notion accounts best for that sort of data. The issue was rather how the single-bottleneck notion can be generally established, at least in the limited sense that queuing is the rule across dual-task situations and concurrent processing is either infeasible or never used.

Since divisible resource models by definition do not preclude queuing, whereas the single-bottleneck notion *does* preclude concurrent processing, it follows that attempting to demonstrate concurrence is more helpful for deciding between the two rival views than demonstrating queuing. Any degree of concurrence should be sufficient to convince one that queuing is not general. On the other hand, if queuing was generally observed even when the circumstances called for concurrent processing, the notion that concurrence is practically prohibited by the existence of a central bottleneck would be selectively supported.

Once the issue is defined to be that of finding out whether concurrent processing of two tasks is a real option, it is easy to realize that the proper tool for doing that is a paradigm that does not induce queuing. The hypothesis that concurrence is an option would be challenged, if it was shown that there

⁹ To illustrate, the difference between predictions from Euclidean geometry and from Riemannian geometry (postulating that no line has any parallel line) for real terrestrial objects is below the resolution of present measuring devices. Their validity can be tested only in the celestial scale. Interestingly, the winner is the nonintuitive theory—Riemannian geometry. One might find a lesson in that: Never disparage small differences.

was no evidence for concurrence even when concurrent processing was called for, namely when the performer had some reason for resorting to concurrent processing that outweighed its disadvantages. Concurrent processing might be preferable either when the tasks require sustained attention (e.g., retention in STM) or periodic monitoring (e.g., tracking) or when heavy time pressure makes queuing fatal to the accuracy of the postponed task (e.g., Gopher, Brickner, & Navon, 1982; Navon, Gopher, Chillag, & Spitz, 1984; Roldan, 1979; Sperling & Melchner, 1978; Tsang, Shaner, & Vidulich, 1995; reviews in Gopher, 1994; Gopher & Donchin, 1986; Wickens, 1984). However, advocates of the single-bottleneck notion have been expending most of their effort in demonstrating queuing in the overlapping tasks paradigm that does *not* call for concurrent processing. One might retrospectively wonder whether that was an optimal choice.

Several dual-task paradigms call for concurrent processing with a varied degree of relative emphasis on the two tasks (Gopher & Donchin, 1986; Gopher & Navon, 1980; Navon & Gopher, 1979, 1980; Wickens, 1984). Despite methodological problems inherent in some of them (see Allport, 1980; Navon, 1984, 1990a), they are at least able to furnish evidence on whether people can adjust task priorities as instructed. It thus seems important to examine how compatible the single-bottleneck is with pervasive phenomena observed in other dual-task paradigms. An extensive discussion falls outside the scope of this article. Suffice it to say that phenomena observed in other dual-task paradigms which were designed to be fairly representative of real world multiple-task situations are not easily reconcilable with the notion that there exists some bottleneck that bars any concurrence whatsoever. Parallel processing of tasks is not ruled out.

Parallel processing does not necessarily entail, of course, that resources are limited and flexibly allocated. Most effects in dual-task performance can be accommodated by a notion of outcome conflict as well (Navon, 1985; Navon & Miller, 1987). That notion seems to be more consonant with neuro-morphic processing than is the notion of divisible resources (Navon, 1998). Kinsbourne (1981) reached a similar conclusion based on some behavioral and neurophysiological findings. On the other hand, it has been claimed that some other facts about the brain are more consonant with resource theory (Gopher, 1986). Since it is hard to decide empirically between those alternatives (Navon, 1984), we chose to pit the Single-Bottleneck Model against a simple model of divisible resources, namely the Single-Resource Model, rather than against some instance of the class of outcome conflict models. In another article (Navon & Miller, 1998) we demonstrate that at least one outcome conflict model can mimic the behavior of the Single-Resource Model, hence fit the data about as successfully as it does.

If indeed concurrent processing was a feasible option, as indicated by those diverse phenomena, findings in the overlapping tasks paradigm would seem of dubious relevance, even if they had been highly diagnostic. One, it

is heavily biased in favor of queuing. Two, it simulates just a small subset of possible cases in which people are called to attain two goals at about the same time. For example, engaging in two continuous tasks, like driving and talking on a cellular phone, seems no less ecologically relevant than making two discrete responses to two stimuli which appear with a temporal gap of a few centiseconds.

The Single-Bottleneck Model and the Parsimony Criterion

One might nonetheless lean toward the Single-Bottleneck Model because it seems most economical. The parsimony criterion is, of course, just a bit of advice about how to model reality, not an a priori postulate about reality itself. As James (1890, p. 453) put it: “. . . Occam’s razor, though a very good rule of method, is certainly no law of Nature.” With no substantial evidence to the contrary, we should opt for the most economical model, but that holds only if the latter is at least equivalent as far as other criteria are concerned. Anyhow, is the Single-Bottleneck Model really the most parsimonious way to account for data about task interference?

As argued above, the Single-Bottleneck Model may actually be viewed as subsumed under resource theory broadly defined. It differs from the Single-Resource Model with respect to assumptions about the divisibility of hypothesized provisions for processing. Yet philosophy of science recommends parsimony in the number of terms and assumptions, not in the number of members of a hypothetical set. A physical theory would not be more economical, if it assumed that the maximal number of electrons in the outer ring of the atom was 1 rather than 8. Furthermore, assuming that that number equals 1 (or any other specific number for that matter) is stronger than just assuming that the number is *some* unknown positive integer. Likewise, parsimony is not compromised by a larger number of tellers in the bank or by a larger number of provisions for processing in the brain. True, that number does have bearing on the tractability of the model. However, tractability is not necessarily a hallmark of a parsimonious model, and it certainly cannot be a serious consideration when validity is concerned.

The notion of divisible resources is thus not inherently uneconomical. Adding to that the assumption of flexible allocation does not seem in itself less parsimonious than assuming instead that allocation was rigid. True, once flexible allocation is allowed, allocation ratio becomes a parameter. However, sparing that parameter requires either assuming that allocation is rigid or assuming that there is nothing to allocate because there is only one provision for processing. Both assumptions are quite strong.

Furthermore, divisible resource models have seldom committed themselves to any ontological assumption about number of processors. Well aware of the difficulty in empirically determining whether processing tasks in a dual-task situation is truly parallel, resource theorists took an agnostic stance and just claimed that processing is concurrent. In other words, they

have been quite willing to accept that the allocation they surmise to occur may actually be allocation of time, namely that it might result from time sharing a central processor (see footnote 1).

True, various resource models went far beyond the basic notion of flexible allocation and took the liberty of proliferating assumptions, e.g., on the number of resource pools or on the shapes and parameters of performance-resource functions. In doing that resource theorists did not always care about parsimony (see critical review in Navon, 1984). On the other hand, all that extra fat was motivated by the need to account for variability in the extent of interference across task pairs and conditions. Resource models do not need that fat to account just for the basic phenomena of task interference. That is taken care of just by a single postulate—that provisions for processing are scarce. The Single-Bottleneck Model adds to that the assumption that the only scarce provision is a single mechanism practically limited to processing one task at a time (at a certain stage).

Furthermore, believers in the single-bottleneck notion often invoke accessory assumptions like response grouping or task preparation to address embarrassments (e.g., Pashler, 1990). One might also add (Pashler, 1994a) modality-specific perceptual processing structures, a single mental timer, or some constraints on joint response production.

There is no reason, of course, why these ancillary assumptions—plausible as they might be—could not be valid if it were also true that a single serial bottleneck is the major limit on dual-task performance. Alas, invoking them does discount the special advantage of the single-bottleneck notion. It is no longer a simple theory that accounts for the data with no crutches. The strict central capacity model was dismissed for less than that (e.g., Allport, 1980; Navon & Gopher, 1979). After all, the favorite escape exits of resource theorists—data limits, elastic capacity, and so on—are just variations on the underlying resource theme. Response grouping, task preparation, and response conflict are, in contrast, foreign imports to the basic thesis of bottleneck models. Although they are typically used to rescue the single-bottleneck notion, they are compatible with other models of task interference as well. Ironically, they can also serve as potential competitors for any model of task interference. For example, task preparation was proposed as the single account for results prevalently interpreted in terms of a single bottleneck (Koch, 1995; but see Ruthruff, Pashler, & Klaassen, 1996).

Worse yet, such assumptions virtually nullify the predictive value of models based on the single-bottleneck notion. If there was any hope to make quantitative predictions, it would greatly wane in view of the need to invoke those additional assumptions. How, for example, can one predict in what proportion of the trials a subject would elect to group her responses?

Thus, parsimony is not the distinctive feature of the Single-Bottleneck Model, let alone of its augmented versions. It may instead be that the real source of appeal of the Single-Bottleneck Model aside from its apparent

tractability is its intuitiveness. The Single-Bottleneck Model may seem more intuitive than does the Single-Resource Model, or for that matter any model that assumes concurrent processing. However, intuitiveness is dubious as a scientific criterion, as the evolution of physics illustrates. Furthermore, the very reason that makes the model intuitive should lead us to examine it with special scrutiny. An intuitive model typically evolves to explain prevalent phenomena. An experimental demonstration that it actually does adds little. Even if the model were wrong, the phenomena that led to its conception are unlikely to indicate that.

One might contend though that intuition about mental processes is more valid since it is based on introspection. The average reader of these words would probably be willing to swear that while he was reading them, his mind could not have accomplished any other goal as well. That introspection, however, has access only to awareness, not to processing which gives rise to what we are aware of. Accordingly, we probably would do well to substitute our prejudice about the feasibility of concurrent processing with what it is worth in hard-core evidence.

A Summary

The present analysis is meant neither as a crusade against the single-bottleneck notion nor as a eulogy of resource theory. It rather asks what is needed to test whether structural queuing is a general phenomenon in dual-task performance.

We argue that evidence from the overlapping tasks paradigm does not necessarily indicate strict queuing because other models can accommodate the data as well. Moreover, even if task interference in that paradigm could be shown to be mediated by queuing, it would not follow that the queuing must be forced by a structural limitation like the existence of a single bottleneck. Simply, the paradigm is inherently biased in favor of queuing.

Examining whether concurrent processing is barred by such a structural limitation requires a different paradigm. The Single-Bottleneck Model does well in accounting for interference between two "simple punctate tasks" (Pashler, 1994a, p. 235), especially for data collected in the overlapping tasks paradigm. However, developing specific models to account for situations that do not fully represent the whole object of study is not quite commendable as a way of procuring parsimony. It certainly does not enhance completeness, another desideratum for scientific modeling. As Logan and Gordon (2001) put it, "The divide-and-conquer strategy that dominates much of current research has clarified different aspects of the blooming, buzzing confusion that surrounds us but it has not yet made clear how the different aspects are related. Viewed collectively, the special-purpose theories may be as confusing as the bloom and the buzz they were meant to clarify" (p. 421). The single-bottleneck notion thus might be too restrictive to provide a general description of dual-task interference.

One might still find the combination of simplicity and intuitiveness in the Single-Bottleneck Model hard to resist, despite the dubious a priori plausibility of the thesis that all customers of that busy establishment sitting in the brain are gated next to the exit by a strictly serial device. But it is about time to weigh that against all other points raised above. Exploring a simple, intuitive thesis like that suggested by the Single-Bottleneck Model surely is a worthwhile scientific investment. But it loses quite a bit of its simplicity and intuitiveness once it needs to be augmented with auxiliary mechanisms such as response grouping, preparation, and so on. With these theoretical fixes, the Single-Bottleneck Model appears practically irrefutable. However, by Popper's (1995) supreme rule, a primary property of any scientific statement is that it is not protected against falsification. Bold attempts to conceive of the world as particularly simple owe science at least that—falsifiability.

APPENDIX A

The Single-Bottleneck Model

This model has already been presented formally elsewhere (e.g., Pashler & Johnston, 1989). We reformalize it here mainly to standardize notation with the other models and to present the basis for our quantitative predictions (see Fig. 6).

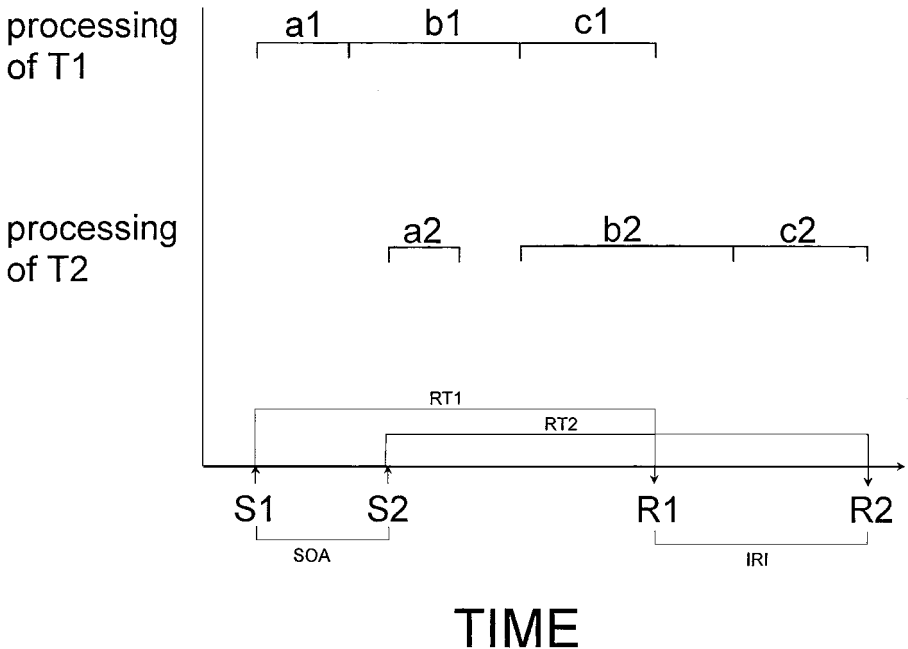


FIG. 6. An illustrative time diagram for the Single-Bottleneck Model.

Let T1, S1, and R1 denote task 1, stimulus 1, and response 1, respectively, and let RT1 be the time from S1 onset to R1 emission. T2, S2, R2, and RT2 are defined analogously for the second task. Let SOA be the time from S1 onset to S2 onset, and let IRI be the time from R1 to R2 (namely $IRI = RT2 - RT1 + SOA$).

According to the Single-Bottleneck Model, processing may be divided into three stages: processing before the bottleneck, processing during the bottleneck itself, and processing after the bottleneck. There is evidence that these stages correspond to encoding, response selection, and response execution processes, respectively, but these labels are not essential to our formalization of the model. We will let a1, b1, and c1 denote the durations of these three processing stages within T1, respectively, and a2, b2, and c2 denote the processing times of the corresponding three stages within T2. We will allow all six of these durations to be random variables, fluctuating from trial to trial within a condition.

The time diagram in Fig. 6 might help in understanding both the assumptions and the following derivations.

The prevalent version of the model assumes that both encoding and response execution can be run in parallel for T1 and T2, but that response selection requires a bottleneck structure that can process only one task at a time. Furthermore, it assumes that T1 is always admitted first to the bottleneck structure. Hence, if T2 were ready for response selection before the response for T1 had been selected, it would have to wait until the bottleneck was free.

It is easy to see that under these assumptions

$$RT1 = a1 + b1 + c1. \tag{1}$$

The response selection processing of T2 either does not have to wait for the bottleneck (a condition termed below: cond1) or has to wait for it (cond2). Hence,

$$RT2 = \max (a1 + b1 - SOA, a2) + b2 + c2, \tag{2}$$

namely

$$RT2 = \begin{cases} a2 + b2 + c2 & \text{if } a1 + b1 - SOA \leq a2 \text{ [cond1]} \end{cases} \tag{3a}$$

$$\begin{cases} a1 + b1 - SOA + b2 + c2 & \text{if } a2 < a1 + b1 - SOA \text{ [cond2]}. \end{cases} \tag{3b}$$

It follows from (3a) and (3b) that

$$E(RT2) = E(b2 + c2) + E(a2|cond1) \cdot \text{prob}(cond1) + E(a1 + b1 - SOA|cond2) \cdot \text{prob}(cond2). \tag{4}$$

Clearly, then, $E(RT1)$ does not depend at all on SOA. On the other hand, $E(RT2)$ decreases with SOA as long as $\text{prob}(\text{cond}2) > 0$. The rate of its decrease increases with $\text{prob}(\text{cond}2)$, reaching an asymptote when $\text{prob}(\text{cond}2) = 1$. In particular, when $\text{prob}(\text{cond}2)$ is practically 1 (namely at the very left of the RT2–SOA curve, at least when encoding S2 is not overly difficult), the slope must be very close to -1 .

IRI. Substituting (1), (3a), and (3b) in the definition of IRI yields:

$$\text{IRI} = \begin{cases} a_2 + b_2 + c_2 - a_1 - b_1 - c_1 + \text{SOA} & : \text{cond}1 & (5a) \\ b_2 + c_2 - c_1 & : \text{cond}2. & (5b) \end{cases}$$

It is straightforward to see that $E(\text{IRI})$ does not depend at all on SOA as long as $\text{prob}(\text{cond}1) = 0$. The slope of the IRI vs SOA curve progressively increases as $\text{prob}(\text{cond}1)$ increases, until it reaches unity at the asymptote where $\text{prob}(\text{cond}1) = 1$. The reason for that increase in slope is that IRI is directly related to SOA under $\text{cond}1$ and is unrelated to SOA under $\text{cond}2$.

Correlation. It is also instructive to consider the various sources of correlation between RT1 and RT2, namely $\text{CORR}(\text{RT}1, \text{RT}2)$, when the stage times (e.g., a_1 or b_2) are random variables. The strength of the correlation depends on the similarity of terms in the equations for RT1 and RT2 [(1), (3a), and (3b)]. Note that when $\text{prob}(\text{cond}1) < 1$, $\text{CORR}(\text{RT}1, \text{RT}2)$ is inflated by variability in a_1 and b_1 but diluted by variability in c_1 , b_2 , and c_2 . In addition, high values of $a_1 + b_1$ and low values of $a_2 + \text{SOA}$ contribute to $\text{prob}(\text{cond}2)$, thereby serving to inflate $\text{CORR}(\text{RT}1, \text{RT}2)$. More important, comparing the terms indicates that $\text{CORR}(\text{RT}1, \text{RT}2)$ should be 0 when $\text{prob}(\text{cond}1) = 1$, assuming that all six stage times are independent. To test this prediction, there should be a way to estimate which level of SOA was sufficiently long to render $\text{prob}(\text{cond}1)$ practically 1. We propose the following rationale.

How long is long? We suggest a disjunctive criterion for determining where $\text{prob}(\text{cond}1)$ is practically 1. That is needed, because RT data are typically affected by nuisance factors and because not all studies report all pertinent data. The criterion states that $\text{prob}(\text{cond}1)$ is practically 1 at an SOA level that satisfies either of the following conditions: (a) the slope of the RT2–SOA curves is very close to 0 or (b) the dual-task effect on the difference between mean RT2 and mean RT1 is negligible.

The first branch of the criterion seems straightforward. As long as $\text{prob}(\text{cond}2)$ is considerable, RT2 must decline with SOA [see Eq. (3b)]. Hence, a negligible slope indicates that the likelihood of $\text{cond}2$ is negligible. That, of course is a conservative criterion, since preparation costs on RT2 (see Appendix E), if such exist, may decline slowly with SOA.

The second branch of the criterion is based on a more complex reasoning.

From the equations for RT1 and RT2 [(1), (3a), and (3b)] it follows that the difference between mean RT2 and mean RT1 at cond1 in the dual-task condition (denoted hereafter as DIF'') equals $a_2 + b_2 + c_2 - a_1 - b_1 - c_1$. The difference between mean RT2 and mean RT1 in the single task condition (denoted hereafter as DIF') has exactly the same form. If there was no dual-task deficit due to preparation cost, cross-talk, and so on, the expression $DIF'' - DIF'$ would be 0. If there were such deficits, the task 1 deficit would probably be smaller than, or at least equal to, the task 2 deficit, hence DIF'' would not be smaller than DIF' . It is, thus, quite safe to conclude that $\text{prob}(\text{cond1})$ must be practically 1 at an SOA level for which $DIF'' - DIF'$ is found to be negligible. Note that this is also a conservative criterion. $DIF'' - DIF'$ may be positive even when $\text{prob}(\text{cond1})$ is about 1 due to two reasons—the effect of preparation cost on task 2 (see above), and the impact on mean RT2 of the small number of trials in which cond2 is met [hence RT2 is larger, since the difference between (3b) and (3a) is positive by the definition of cond2].

Arousal effects. There may be trial-to-trial fluctuations—in arousal or any other nuisance factor—that have a common influence on all terms, leading to a source of correlation beyond that inherent in the model. There are any number of ways to augment the Single-Bottleneck Model to allow for the possibility of arousal fluctuations and other trial-to-trial effects that could produce correlations between RT1 and RT2 even if cond1 were satisfied on every trial. For example, one simple possibility is to assume trial-to-trial variation in the rate of bottleneck processing. Let the rate be denoted by q ($0 < q \leq 1$), where q is sampled randomly for each trial. The coefficient $1/q$ then multiplies the b_1 and b_2 terms in all the above expressions for a given trial. The most interesting consequence of this type of variability is that $\text{CORR}(\text{RT1}, \text{RT2})$ is no longer 0 when $\text{prob}(\text{cond1}) = 1$ because $\text{CORR}(\text{RT1}, \text{RT2})$ is inflated by variability in q .

APPENDIX B

The Single-Resource Model

Single-Resource Models assume that processes are nourished by resources from a single finite pool. To simplify the model we use and maximize its similarity to the Single-Bottleneck Model, it is further assumed that for any given task, the only stage of processing that is resource-limited is the response selection stage—the second of three processing stages.

One of the defining characteristics of resource models is that resources are allocated between processes which make concurrent demands (Navon, 1985). In the context of the overlapping tasks paradigm, that entails that resources are allotted to the response selection stage of the second task as soon as the encoding of S2 is terminated.

Resources are assumed to control the rate of processing by a resource-limited stage, with the allocation dividing the total rate between the two tasks. The total time needed by the resource-limited stage equals its demand divided by the rate allocated to it. As in the simulations of the Single-Bottle-neck Model, we also allowed trial-to-trial variations in arousal to modulate the total processing rate which was divided between the two tasks.

Let us use the same notation used in Appendix A with the following changes: d_1 and d_2 are demands for resources by response selection processing for T1 and T2, respectively. Dividing each by processing rate yields response selection processing times, b_1 and b_2 . Let the single-task rate under maximal arousal be set to 1, with no loss of generality. Let the actual single-task rate under submaximal arousal be denoted by q ($0 < q < 1$). Let the proportion allocated to the first task be denoted by p ($.5 < p < .9$).

Figure 7 illustrates how in the Single-Resource Model the two tasks in the overlapping tasks paradigm might use resources.

The response selection for T1 can use all available resources until T2 is ready for response selection (during Only-T1-Period, or OT1P for short). It might be terminated by then (cond1). Otherwise, namely if by then R1 has not yet been selected (cond2), T1 processing could subsequently use only a proportion p of available resources (during the OVerlap-Period, or OVP for short). For simplicity, we exclude the possibility that selection of R2 finishes before selection of R1 because this possibility is remote with the particular tasks ordinarily used in studies with the overlapping tasks paradigm. Thus,

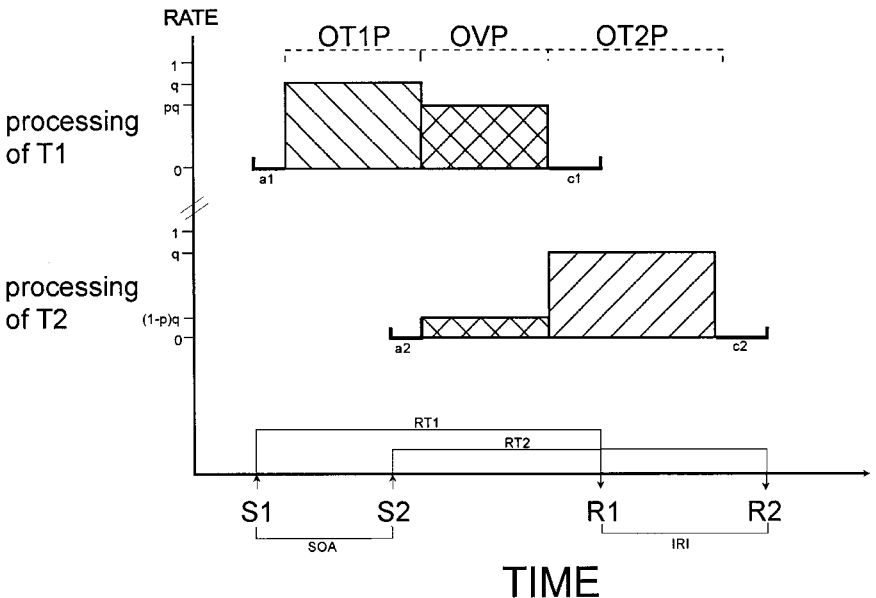


FIG. 7. An illustrative time diagram for the Single-Resource Model.

$$RT1 = a1 + b1 + c1, \tag{1}$$

namely

$$\begin{aligned}
 RT1 = & \\
 a1 + c1 + & \left[\begin{array}{l}
 d1/q \quad : d1 \leq q(SOA + a2 - a1) \text{ [cond1]} \\
 \left| \begin{array}{l}
 OT1P + [(d1 - \text{work done at OT1P}) / (p \cdot q)] = \\
 (SOA + a2 - a1) + \\
 \left| [d1 - q(SOA + a2 - a1)] / (p \cdot q) : q(SOA + a2 - a1) < d1 \text{ [cond2]}. \end{array} \right. \\
 \left. \right. \end{array} \right. \tag{2a}
 \end{aligned}$$

The response selection for T2 can use all available resources, if by its start R1 has already been selected (namely under cond1). Otherwise (under cond2), it must start using only a proportion $1 - p$ of available resources. That lasts as long as selecting the response for T1 has not yet terminated (namely during OVP). Once R1 has been selected, T2 can use all available resources for selecting R2 (during the Only-T2-Period, or OT2P for short). Thus,

$$RT2 = a2 + b2 + c2, \tag{3}$$

namely

$$\begin{aligned}
 RT2 = & \\
 a2 + c2 + & \left[\begin{array}{l}
 d2/q \quad : \text{cond1} \\
 \left| \begin{array}{l}
 OVP + (d2 - \text{work done at OVP}) / q = \\
 [d1 - q(SOA + a2 - a1)] / (p \cdot q) + \\
 \left| \{d2 - \{[d1 - q(SOA + a2 - a1)] / (p \cdot q)\} \cdot (1 - p) \cdot q\} / q : \text{cond2} \end{array} \right. \\
 \left. \right. \end{array} \right. \tag{4a}
 \end{aligned}$$

It follows from (2a)–(2b) that

$$\begin{aligned}
 E(RT1) = & E(c1) + E(a1 + d1/q | \text{cond1}) \cdot \text{prob}(\text{cond1}) \\
 & + E[a1/p + d1/(p \cdot q) \\
 & + (1 - 1/p) \cdot (SOA + a2) | \text{cond2}] \cdot \text{prob}(\text{cond2}) \tag{5}
 \end{aligned}$$

It follows from (4a)–(4b) that

$$E(RT2) = E(c2) + E(a2 + d2/q | cond1) \cdot \text{prob}(cond1) \\ + E[(d2 + d1)/q + a1 - SOA | cond2] \cdot \text{prob}(cond2). \quad (6)$$

Thus, $E(RT1)$ does not depend on SOA, when $\text{prob}(cond1) = 1$, namely when SOA is particularly large or $d1$ is particularly small. Otherwise, $E(RT1)$ decreases with SOA with a rate of decrease that is inversely related to $\text{prob}(cond1)$. However, even when $\text{prob}(cond1) = 0$, the slope depends on p . For example, when $p = .5$, the slope is -1 ; when $p = .9$, the slope is $-.11$. Interestingly, under the latter condition, the dependence of $E(RT1)$ on $a1$ and $d1$ is modulated by p : The smaller p is, the higher the impact of $a1$ and $d1$.

$E(RT2)$ decreases with SOA as long as $\text{prob}(cond1) < 1$, namely when SOA is not particularly large and $d1$ is not particularly low. The rate of its decrease is inversely related with $\text{prob}(cond1)$, until it reaches an asymptote when that probability equals 1. In particular, when $\text{prob}(cond2)$ is practically 1 (namely at the very left of the RT2–SOA curve), the slope must be very close to unity. Interestingly, under the latter condition, $E(RT2)$ does not depend on $a2$ at all.

Dual-task decrement. For considering dual-task decrements, some simplifying assumptions might help the exposition. Suppose $a1 = 0$, $a2 = 0$, $c1 = 0$, $c2 = 0$, $q = 1$, and $d1 = d2 = d$ (hence $b1 = b2$). Further suppose that $p(\text{cond2}) = 1$. It follows from (1) and (3) that in the single-task conditions, $RT1 = RT2 = d$. Latencies in the dual-task condition depend, however, on p , namely the proportion of resources allotted to task 1. It follows from (5) and (6) that $RT1 = RT2 = 2 \cdot d - SOA$ when $p = .5$, thus both tasks suffer the same dual-task decrement. However, when $p = .8$, for example, $RT1 = 1.25 \cdot d - .25 \cdot SOA$ but $RT2 = 2 \cdot d - SOA$. Hence, the dual-task decrements equal $d - SOA$ for both tasks when $p = .5$, but $.25 \cdot (d - SOA)$ and $d - SOA$, respectively, when $p = .8$. In other words, the dual-task decrement for RT2 is unaffected by allocation ratio, but the dual-task decrement for RT1 is smaller as the priority of task 1 is increased.

IRI. Substituting (2a)–(2b) and (4a)–(4b) in the definition of IRI yields:

$$IRI = \begin{cases} a2 + d2/q + c2 - a1 - d1/q - c1 + SOA & :cond1 \end{cases} \quad (7a)$$

$$\begin{cases} d2/q + [(p - 1)/(p \cdot q)] \cdot d1 + \\ c2 - (1 - 1/p) \cdot (a2 + SOA - a1) - c1 & :cond2. \end{cases} \quad (7b)$$

Thus, unlike in the Single-Bottleneck Model, here $E[IRI]$ must depend on SOA. As long as $\text{prob}(cond2) = 1$, the slope is moderate provided that p is

sufficiently greater than .5. Then the slope of the IRI–SOA curve progressively increases as $\text{prob}(\text{cond}2)$ decreases. When $\text{prob}(\text{cond}1) = 1$, it reaches unity.

Correlation. Comparing (2a)–(2b) with (4a)–(4b) suggests that $\text{CORR}(\text{RT}1, \text{RT}2)$ must almost always be positive. Under $\text{cond}2$, it is inflated by variability in a_1 , d_1 , and q , but diluted by variability in c_1 , a_2 , b_2 , and c_2 . Under $\text{cond}1$, $\text{CORR}(\text{RT}1, \text{RT}2)$ is inflated by q and diluted by all other factors.

Like in the Single-Bottleneck Model, low values of $a_2 + \text{SOA}$ reduce $\text{prob}(\text{cond}1)$, thereby serving to inflate $\text{CORR}(\text{RT}1, \text{RT}2)$. Since $\text{prob}(\text{cond}1)$ grows as SOA is increased, the correlation would decrease. However, unlike in the Single-Bottleneck Model, the correlation would still be positive due to variability in q .

Arousal effects. Here the assumption about arousal is inherent in the definitions. The limit on available resources, q , has been assumed to be directly related to level of arousal.

APPENDIX C

Simulation Programs

This appendix describes the methods by which we obtained the illustrative fits of the Single-Bottleneck Model and the Single-Resource Model to the mean RTs of various data sets, as shown in Figs. 1–5.

Of course, the exact predictions of the models depend on details of the assumptions made in implementing the models. Accordingly, our goal was to show just that there exists at least one version of each model that is capable of reproducing the main features of the data. That turned out to be rather easy to do.

Each of the models has been presented formally in its own appendix, but a comparison across models shows that they have similar structures and many parameters in common. In particular, the parameters a_1 , a_2 , c_1 , and c_2 enter into both models in approximately the same way, and so we decided to handle these parameters in a uniform fashion across both models. For this reason and others, it is most convenient to describe in detail how one model was fit and then describe the modifications that were needed to fit the other one. It is simplest to start with the Single-Bottleneck Model.

The Single-Bottleneck Model

Within the Single-Bottleneck Model, the results for any trial (i.e., RT1, RT2, and IRI) are uniquely determined by the values of a_1 , b_1 , c_1 , a_2 , b_2 , c_2 , and SOA for that trial. A quantitative specification of the Single-Bottleneck Model, then, requires a specification of how all of these values vary and covary from trial to trial. It is most realistic to assume that all six values are

random variables. This leads to a computationally complex model, however, and so we adopted the simplifying assumption that the values of a_1 , a_2 , c_1 , and c_2 were constants, as described further below. This assumption is probably an acceptable working approximation if the variances of these values are small relative to those of b_1 and b_2 , as seems plausible.

We arbitrarily set a_1 and a_2 to be constants of 50 ms in all conditions except the low intensity task 2 condition of Pashler and Johnston (1989, Fig. 3), where we set a_2 to be a constant of 85 ms. The 35-ms difference (i.e., 85 vs 50) was suggested by the maximal intensity effect obtained at the longest SOAs.

The values of c_1 and c_2 were also assumed to be constants, both across trials and across conditions, although they were allowed to differ from one another. The values of these constants were not chosen arbitrarily, however, but instead were adjusted to minimize the deviations between the predicted and observed values, subject to the limitation that both fall within the range of 50–300 ms.

Both b_1 and b_2 were assumed to be independent random variables, each drawn from its own probability distribution. These probability distributions were modeled as gamma distributions because the gamma is a theoretically plausible model of finishing times (Luce, 1986; McGill, 1963). Each gamma was assumed to have a shape parameter of 3 to produce a moderate positive skew as is common in RT distributions, and the rate of each gamma distribution was estimated by a numerical parameter search process, as described further below. To simplify the computations, we approximated each distribution using a discrete 10-point approximation consisting of 10 equally likely values from the 5th, 15th, . . . , 95th percentiles of each gamma distribution.

Most of the data sets included several experimental conditions, and the values of b_1 or b_2 were expected to vary across conditions. The effect of condition was always modeled as a constant, with the values of the constants estimated by eye for each data set separately. For example, the data set of McCann and Johnston (1992, Fig. 3) includes ordered and arbitrary conditions in which the values of b_2 are expected to vary. For this data set, we modeled the values of b_2 in the ordered condition with 10 percentile points as described above and then increased each of these 10 percentile points by 62 ms in the arbitrary condition, where 62 ms was just an eyeball estimate of the overall ordered arbitrary difference in this data set. Analogously, for the data set of Pashler and Johnston (1989, Fig. 4), nonrepetitions were modeled as having b_2 values 25 ms larger than repetitions. For the data of Smith (1969), the values of b_1 in the four- and eight-alternative tasks were modeled as being 120 and 190 ms larger, respectively, than those in the two-alternative task.

In summary, the model had four free parameters: c_1 , c_2 , g_1 , and g_2 , where g_1 and g_2 are the means of the gamma distributions from which b_1 and b_2 were derived (at the fastest level of any factor affecting these rates). The parameters g_1 and g_2 were iteratively adjusted using the simplex algorithm

(Rosenbrock, 1960) to find values that would minimize the mean-squared error between predicted and observed mean RT1s and RT2s.

For each pair of g_1 and g_2 values that were examined, the optimal values of c_1 and c_2 were determined algebraically, as described next. Given a candidate pair of g_1 and g_2 values, the program proceeded as follows. First, c_1 and c_2 were arbitrarily set to 100 ms. Second, the program simulated 100 trials in each of the conditions of the experiment, using all 10×10 combinations of different b_1 and b_2 values corresponding to the gamma distributions derived from those g_1 and g_2 values (and increased by constants representing factor effects, as necessary). The predicted RT1 and RT2 were computed for each trial in each condition. Third, the results from the 100 simulated trials in each condition were averaged to determine the mean predicted RT1s and RT2 for that condition. Fourth, the values of c_1 and c_2 were adjusted to optimize the fit of the model with the current g_1 and g_2 values, subject to the constraints that they be within the range of 50–300 ms. To do this, the overall average predicted RT1 was determined, averaging across conditions, and compared with the overall average observed RT1. The value of c_1 was adjusted so that these overall averages would be exactly equal, and all of the individual predicted RT1s were adjusted accordingly. For example, if the predicted average RT1 were 400 but the observed average RT1 were 450, then the value of c_1 would be increased from 100 to 150, which would increase the predicted RT1 by 50 ms in every condition. The value of c_2 and predicted RT2s would similarly be adjusted to match predicted and observed RT2s on average across conditions. Fifth, the mean-squared error of the predicted versus observed RT1s and predicted and observed RT2 was computed, and this was the summary measure of the goodness of fit for the given candidate g_1 and g_2 values.

Table 1 summarizes the results of the parameter-fitting process. For each data set, the table shows the estimated value of each free parameter, plus an overall measure of the quality of the fit (root-mean-square error, RMSe, in milliseconds).

Note, however, that parameter values are not identifiable in the sense that different combinations of parameter values can yield identical predictions about the observable RTs. For example, reductions in b_2 can be compensated completely by increases in c_2 because the model's predictions depend only on the total $b_2 + c_2$, not on the division of this total into its two components. The implication of this nonidentifiability is that parameter estimates are by no means unique, which in turn makes it difficult to decide whether the model can fit with "reasonable" parameter values. The same is true to some extent of the other model as well.

The Single-Resource Model

The experimental task demands and condition effects were modeled in the same way as for the Single-Bottleneck Model, but there is a small notational difference. In this case we modeled d_1 and d_2 instead of b_1 and b_2 . The

terms d_1 and d_2 denote the demands of the two tasks on the response selection stage, namely bits of information to be transmitted or units of resources needed to transmit them. Operationally, these may be measured in terms of the number of milliseconds that each task would require if it had 100% of resources allocated, so they correspond conceptually to the parameters b_1 and b_2 of the Single-Bottleneck Model. However, the actual processing times b_1 and b_2 depend not only on d_1 and d_2 but also on the amount of overlap of the two processing stages and on the relative priorities given to the two tasks, as shown in Eqs. (2) and (4) of Appendix B. Thus, we modeled d_1 and d_2 using gamma distributions, just as we had done for the analogous b_1 and b_2 in the Single-Bottleneck Model, and simply computed b_1 and b_2 from d_1 , d_2 , and the other necessary parameters of the Single-Resource Model.

As for p , namely the proportion of resources allocated to the first task, the only constraint on it in the model is that it is <1 . However, to make the tested models sufficiently different, we set p to .9 in the simulations. This parameter was assumed to be constant across all conditions and trials within a simulation. We also ran a simulation with p running from .50 to .95 only on the data of the single experiment in which the instructions were not explicitly biased (Pashler, 1991, Fig. 8).

The value of the arousal parameter q was fixed at 1 for these simulations. This parameter seems important primarily for considering why RT1 and RT2 might be positively correlated, but less for considering the mean values of RT1 and RT2 that might be predicted.

Parameter estimation followed the same basic procedure used for the Single-Bottleneck Model. Given a candidate set of parameter values g_1 and g_2 , the program simulated 100 trials in each condition of the experiment, using all 10×10 combinations of different d_1 and d_2 values. The predicted mean RT1 and RT2 were computed for each condition, and c_1 and c_2 were adjusted to optimize the fit across all conditions. Table 1 shows the estimated parameters and goodness-of-fit measures for each of the data sets.

APPENDIX D

Can Response Grouping Help the Single-Bottleneck Model Much?

The notion of response grouping is needed primarily to reconcile the Single-Bottleneck Model with effects that appear to support concurrent processing (see the section "Other Relevant Findings"). In addition, however, it has often been used to explain aspects of data obtained in the overlapping tasks paradigm that were incompatible with the Single-Bottleneck Model. Assuming that subjects sometimes group their responses in studies using that paradigm actually generates a new model, the augmented Single-Bottleneck Model, that must be evaluated using the same criteria as the others.

That people can group their responses is a fact. That subjects sometimes

do it is probably true as well. The issue is whether they do it frequently enough to render response grouping a noteworthy determinant of performance in the overlapping tasks paradigm. Note that to account for those RT1 effects which the Single-Bottleneck Model cannot accommodate, the proportion of trials on which subjects group their responses must be considerable, sometimes up to .5 (e.g., Pashler, 1984; also Smith, 1969).

The traditional assumption was that response grouping occurs only when SOA is particularly short (Welford, 1967). The sense that may be worked into this originally post hoc assumption is that subjects use a grouping strategy to avoid the cost of a refractory effect of R1 and that such a cost is large only when SOA is particularly short. However, present theorizing of the Single-Bottleneck Model does not make this restriction (e.g., Pashler & Johnston, 1989), at least not a priori.

Nonetheless, a model may be valid even when its functional basis is unclear. It is thus interesting to examine what would happen had we let the Single-Bottleneck Model use this auxiliary assumption. For that purpose we derived precise predictions from a version of the Single-Bottleneck Model augmented by response grouping. We used a logic proposed by Pashler and Johnston (1989). The major change in the model is that a strategy of response grouping is assumed to be used, possibly with a probability that declines with SOA. The analysis is presented below, followed by a discussion of its implications for the support of the Single-Bottleneck Model.

The Single-Bottleneck Model Augmented by Grouping

This model is the same as the one presented in Appendix A, except that a strategy of response grouping is assumed to be used on a given trial with probability r . It has been suggested by early proponents of the Single-Bottleneck Model that r declines with SOA. We present derivations with and without this auxiliary assumption of a negative effect of SOA on r .

When the strategy is applied, the two responses are emitted by a single process that orchestrates them to be given in close succession. That process starts as soon as both responses have been selected, but not earlier. The processing time of that grouped responses process surely varies, but the interval it allows between the responses must be very short with a very small variability.

Let us use the same notation used in Appendix A, adding just the terms $c'1$ and $c'2$ for denoting the times it takes the grouped responses process to produce R1 and R2, respectively. Actually, it seems reasonable to assume that the processing system tries to control the interval between responses, so that the orthogonal factors are $c'1$ and that interval ($c'2 - c'1$, in the above notation). Nevertheless, we chose to refer rather to $c'2$ as an independent factor to conform to the notation used by Pashler and Johnston (1989).

The equations for RT1 and RT2 in the trials with no response grouping are the same as in Appendix A. In a trial with grouped responses, however,

$$RT1 = \max (a1 + b1, SOA + a2) + b2 + c'1, \quad (1)$$

namely

$$\left[SOA + a2 + b2 + c'1 \quad : a1 + b1 - SOA \leq a2 \text{ [cond1]} \right. \quad (2a)$$

$$\left. \left[a1 + b1 + b2 + c'1 \quad : a2 < a1 + b1 - SOA \text{ [cond2]}. \right. \right. \quad (2b)$$

Also, in a trial with grouped responses,

$$RT2 = \max (a1 + b1, SOA + a2) + b2 + c'2 - SOA, \quad (3)$$

namely

$$a2 + b2 + c'2 \quad : a1 + b1 - SOA \leq a2 \text{ [cond1]} \quad (4a)$$

$$a1 + b1 - SOA + b2 + c'2 \quad : a2 < a1 + b1 - SOA \text{ [cond2]}. \quad (4b)$$

It follows that

$$\begin{aligned} E(RT1) = & E[r \cdot (SOA + a2 + b2 + c'1) + (1 - r) \cdot (a1 + b1 + c1) \mid \text{cond1}] \\ & \cdot \text{prob}(\text{cond1}) + E[a1 + b1 + r \cdot (b2 + c'1) + (1 - r) \\ & \cdot c1 \mid \text{cond2}] \cdot \text{prob}(\text{cond2}) \end{aligned} \quad (5)$$

and that

$$\begin{aligned} E(RT2) = & E[a2 + b2 + r \cdot c'2 + (1 - r) \cdot c2 \mid \text{cond1}] \\ & \cdot \text{prob}(\text{cond1}) + E[a1 + b1 - SOA + b2 \\ & + r \cdot c'2 + (1 - r) \cdot c2 \mid \text{cond2}] \cdot \text{prob}(\text{cond2}). \end{aligned} \quad (6)$$

Thus, trials with grouped responses do not introduce a drastic change in the predictions with respect to RT2, but they lengthen RT1 and make it more dependent on b2 and to a lesser extent on a2. On the other hand, they make RT1 somewhat *grow* with SOA.

Note, however, that if grouping was less frequent at higher SOAs, then all the above effects of grouping would be expected to interact with SOA. That would especially moderate the positive effect of SOA due to the rise in the probability of cond1 because if r declined with SOA, then it would be fairly low when cond1 was satisfied (namely when $a1 + b1 - SOA \leq a2$). If the decline in the frequency of grouping were fast and large enough,

it could even outweigh that positive effect so that at its left the RT1–SOA curve would even have a negative slope.

IRI. It follows from Eqs. (1) and (3) that in trials with grouped responses

$$IRI = c'2 - c'1, \tag{7}$$

hence,

$$\begin{aligned} E(IRI) = & E[r \cdot (c'2 - c'1) + (1 - r) \cdot (a2 + b2 + c2 - a1 - b1 - c1 + SOA) \mid \text{cond1}] \\ & \cdot \text{prob}(\text{cond1}) + E(r \cdot (c'2 - c'1) \\ & + (1 - r) \cdot (b2 + c2 - c1) \mid \text{cond2}) \cdot \text{prob}(\text{cond2}). \end{aligned} \tag{8}$$

Thus, just like in the unaugmented version of the Single-Bottleneck Model, $E(IRI)$ does not depend at all on SOA as long as $\text{prob}(\text{cond1}) = 0$. Then the slope of the IRI–SOA curve progressively increases as $\text{prob}(\text{cond1})$ increases, until it reaches its maximum at the asymptote where $\text{prob}(\text{cond1}) = 1$. However, the maximum may not be unity, if $0 < r \mid \text{cond1}$.

On the other hand, this model may accommodate a faster, and perhaps earlier, increase of the IRI–SOA curve, if it is assumed that r declines with SOA. In that case, as SOA is increased, the proportion of trials with grouped responses decreases. Since IRIs are extremely short in trials with grouped responses, the mean of IRI within a given condition (namely cond1 or cond2) thus increases with SOA.

Correlation. Comparing Eq. (1) with Eq. (3) suggests that in trials with grouped responses $\text{CORR}(RT1, RT2)$ must be very high. Hence, those trials inflate the overall correlation to the extent that they are present. If grouping were equally likely across all SOA values, that would lead to $0 < \text{CORR}(RT1, RT2)$ even when $\text{prob}(\text{cond1}) = 1$, unlike in the unaugmented version of the Single-Bottleneck Model. However, considering that r may decline with SOA, it probably is very close to 0 when $\text{prob}(\text{cond1}) = 1$. Thus, this model is still strained in accounting for positive RT1–RT2 correlations at high SOA values.

What Do We Gain from the Augmented Model?

The augmented model clearly cannot add much to the fit of the sets of data considered here, since those focus mainly on effects that are least affected by response grouping. Indeed, the fit of the Single-Bottleneck Model was good, which suggests that there was little grouping in the experiments examined. The main improvement of the augmented version above the unaugmented one is that it can better accommodate the sensitivity of RT1 to factors related with the presence or difficulty of the second task. With the additional assumption that grouping is more frequent at shorter SOAs, it can accommodate as well the increase in IRI with SOA. However, it is still strained to account for cases in which RT1 decreases with SOA. It would have to invoke

a very fast and substantial decline in the probability of grouping with SOA to account for that.

Thus, at the cost of parsimony, the already-augmented Single-Bottleneck Model can be further augmented to account for any effect. However, to make any sense the impact of grouping must be consistent across effects. If, for example, grouping effectuates a dual-task decrement for task 1 that is half the dual-task decrement for task 2 (Pashler, 1984), it is expected to effectuate also an effect of task 2 difficulty on RT1 that is half its effect on RT2. We do not know of any study that afforded such a comparison.

But more important, does the augmented version of the Single-Bottleneck Model really model the behavior of a subject? Note that response grouping could be a strategy that several subjects always employ while other subjects do not use at all. The augmented version of the Single-Bottleneck Model would then appear to fit the data thanks to averaging across those types of subjects. Furthermore, discarding the groupers would enable direct testing of the unaugmented version of the Single-Bottleneck Model. Thus, the augmented version of the Single-Bottleneck Model would be of interest only if most subjects do group their responses on a significant proportion of the trials (as often conjectured, e.g., Pashler, 1984; Pashler & Johnston, 1989). The hypothesis that they do may be labeled the response grouping mixture hypothesis. Is there any support for it? We do not know of any systematic effort to test the grouping mixture hypothesis, and we wonder whether if such a test was attempted, it would turn out favorable to the hypothesis.

The reason is the following: According to the mixture hypothesis, IRIs may arise from either of two sources. Grouping entails very small IRIs, since with both responses ready and no motor conflict, it would be natural to emit R1 and R2 almost simultaneously. Indeed, Pashler (1994b) found that some subjects had most of their IRIs much smaller than 100 ms and sometimes smaller than 10 ms. For those subjects the variance of IRI was very small and the shape of the IRI distribution was like a "spike." It is safe to assume that those subjects grouped their responses on most trials. This was observed under complete task symmetry, so perhaps the distribution is somewhat elevated in conditions of task asymmetry. However, its variance must still be very small, since it is probably caused only by the unreliability of a motor program that aims at producing a fixed interval between the responses, not by variability in processing time. Some idea about the order of magnitude of that variance is given by a simple-RT experiment reported by Ulrich and Stapf (1985). They asked subjects to respond with both hands to the same stimulus and obtained a range of IRI standard deviations between 6 and 17 ms (and a range of means between 1 and 8 ms). The characteristics of the IRI distribution must be completely different when no response grouping is attempted: Its mean must be high, since even with a small SOA, it contains two positive terms and only one negative term [Appendix A, Eq. (5b)]. More importantly, its variance must be high, since it includes variability in pro-

cessing times. When there are no fluctuations in arousal, it is a variance of uncorrelated differences ($c_2 - c_1$) plus another term (b_2), hence a sum of three variance components. If the findings of Pashler (1994b) are any indication, the IRI distributions of subjects who did not seem to group at all had means and standard deviations that appear several times larger than the corresponding parameters of "spike" IRI distributions. Because of this large disparity of IRI distributions obtained under the two strategies, the distribution of IRIs of subjects who mix both with a medium grouping probability is most likely either bimodal or at least skewed in a way that is not predicted by any reasonable density function. It seems that the onus of proof is on advocates of the mixture hypothesis to demonstrate that that is indeed the case, since parsimony favors any of the pure-strategy hypotheses.

Granted, the Single-Bottleneck Model proponents might suggest a different conception of the strategy subjects use on trials in which they do not emit R1 as soon as it is ready. They might contend, for example, that the strategy has little effect on the parameters of the IRI distribution, at least not in the particular experiments examined here. That actually might boil down to a "strategy" of spacing the responses on a random basis with a single constraint: that $RT_1 < RT_2$. As a variant of the response grouping hypothesis, that seems implausible given the findings of Pashler (1994b) as well as the considerations discussed above. As just another hypothesis, it is too unconstrained to qualify as an object of study. It seems more like a speculation about the source of noise in the data. Nevertheless, such a possibility cannot be ruled out easily. However, does it have to be ruled out? If the onus of proof for the validity of the mixture hypothesis lies with its advocates, that is even more true of this uncommitted version of the hypothesis. Invoking a strategy that generates an IRI distribution with indefinite parameters does not constitute such a proof. Though an account that is not particularly specific is not unheard of in psychology, an account that is meant as an ancillary assumption of a rigorous model should rather be specific. If it was rather generously flexible, then it could be used as a trump card drawn whenever the model seems to be in trouble. However convenient, that is a sure recipe for unfalsifiability. This is especially true of the Single-Bottleneck Model augmented by the uncommitted mixture hypothesis. Since the response grouping account predicts just the opposite of what the Single-Bottleneck Model predicts with respect to the most diagnostic aspect of the data, namely effects on RT_1 , the unspecific mixture can never fail in postdiction, just like a theory claiming that all people are good except for the few who happen to be bad.

Finally, does response grouping suggest much about the source of task interference? If the winner in a horse race was magnanimous enough to wait for the slower horse before crossing the finish line, that still would not explain why the slower horse could have done better had it run alone. Neither is that effect predicted by such courteous conduct. A strategy of grouping

could be used to coordinate response timing even if tasks were processed in parallel with no interference whatsoever. Furthermore, such a strategy could be used, even if interference were due to any of the alternative causes mentioned above or, for that matter, to any other conceivable cause. Thus, response grouping neither explains task interference nor selectively supports the Single-Bottleneck Model.

Accordingly, if subjects indeed tended to frequently group responses, the implication would not be theoretical but rather methodological: Grouping would better be regarded as a nuisance factor in the study of task interference, much as saccadic eye movements are in visual attention experiments. Subjects should be dissuaded from grouping by means of procedure and instructions. In addition, data of subjects suspected of grouping should be trimmed off after data collection, if a reasonable procedure for diagnosing that could be found. Hopefully, that restriction would not thin out the set of data too much. Judging by the arguments we cited above, it would not.

APPENDIX E

Is the Notion of Task Preparation More Helpful?

Another auxiliary assumption often used by the Single-Bottleneck Model proponents appeals to the notion that some preparation is needed before processing starts (Bertelson, 1967; Gottsdanker, 1979, 1980; see reviews in Pashler, 1994a, 1994c, 1998).

If preparing two different S-R mappings takes more time than preparing just one, that might explain why RT1 is considerably affected by the dual-task requirement when there is no evidence for response grouping (Pashler, 1984). A similar idea could be used to explain away the presence of effects on RT2 when SOA is much longer than RT1 (Davis, 1957; Kafry, 1971; Rabbitt, 1969; Triggs, 1968; Welford, 1959). Since in that case processing of S2 is not delayed by having to queue for an engaged bottleneck, the Single-Bottleneck Model proponents have to resort to another account. They might find it helpful to assume that the processing of task 1 must inhibit preparation for task 2 (see Pashler, 1994a)—an even more obvious instance of outcome conflict.

The issue is not whether task preparation effects exist—they do—but rather how their possible existence bears on the validity of the Single-Bottleneck Model as a major model of task interference.

First, is task preparation required or potentially useful *only* when there is a structural bottleneck? Not at all. If the tasks could be processed in parallel, an interference of the sort discussed above might be regarded as increasing the demand of selecting the response for task 2, as decreasing the rate of doing so, or as forcing a preparatory operation of adjusting the process of selecting the response for task 2 to the aggravating circumstances. If the cost

were still too high, it might be averted by strategic queuing. If instead the tasks had to queue at some point for structural reasons, such an interference might again be regarded as increasing the demand of selecting the response for task 2, as decreasing the rate of doing so, or as forcing a preparatory operation of adjusting the process of selecting the response for task 2 to the aggravating circumstances. Advocates of the Single-Bottleneck Model seem to favor the latter possibility, namely that the processing of task 1 adds, or prolongs, an intermediary stage of preparing the response selection process for task 2 (Gottsdanker, 1979, 1980; Pashler, 1994a). It is understandable why from their perspective they prefer that possibility. It is, however, clear that an interference of task 1 with task 2 which is not accounted for by postponement is compatible with any model we have considered so far. Thus, the notion of task preparation could augment any other account of task interference, the Single-Resource Model for one.

Furthermore, that the notion of preparation is often favored by proponents of the Single-Bottleneck Model as an accessory assumption might make us gloss over a simple truth: That notion is actually a rival of the notion of structurally induced postponement.

Note that task preparation effects are limited to cases in which the overlapping tasks differ, at least in their S-R mappings, because if two tasks are identical, then the same preparation would serve for both. Furthermore, note that the magnitude of the effects is often assumed to be related to task similarity. For example, the S-R mapping of task 1 might interfere with setting or implementing the S-R mapping for task 2 to the extent that those are incongruous, confusable, or mutually inhibitory. Preparation is needed to avert that interference.

Interference whose magnitude is task-pair-specific is clearly a variant of outcome conflict (e.g., Navon, 1985; Navon & Miller, 1987). How about interference with a constant magnitude that is nonetheless conditional on some relationship between tasks, namely that they are not identical? That contingency suggests that the interference must be due to some disparity in preconditions. Without preparation, the second task would operate in suboptimal preconditions that have been set to optimize the first task. That is, an instance of interference that is not due to competition for some provision for processing but rather to the adverse effect resulting from concurrent processing—namely an outcome conflict. Thus, in that case as well, task preparation is used to avert outcome conflict.

The notion of outcome conflict is clearly a theoretical rival of the notion of a single bottleneck. By positing that concurrent processing is a feasible option that is sometimes averted to reduce outcome conflict, it entails that the reason for serial processing is not that there is only one processor. Task preparation is a special version of the general notion of outcome conflict: It assumes an adverse effect that is so slow to dissipate that queuing is not enough to minimize it, so that further preparation is called for.

If task 1 prolonged the preparation for task 2, that could in theory be the most potent source of interference (Koch, 1995), even if a structural bottleneck existed. To illustrate, bottleneck processing could last only a few centi-seconds, whereas recovering from the outcome conflict could last many times longer. Thus, the combination of the Single-Bottleneck Model with the notion of task preparation is actually a marriage between rival theories.

Thus, some proactive task interference may be very reasonably assumed. It might or might not affect preparation for task 2. Whichever the case may be, the notion of structurally induced queuing is not selectively supported by that. Furthermore, if the Single-Bottleneck Model—augmented or unaugmented—can fit the data thanks only to assuming a very high degree of proactive interference, then structurally induced queuing can hardly be claimed to be the major determinant of task interference.

REFERENCES

- Allport, D. A. (1980). Attention and performance. In G. Glaxton (Ed.), *Cognitive psychology: New directions* (pp. 112–153). London: Routledge & 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. (1966). Central intermittency twenty years later. *Quarterly Journal of Experimental Psychology*, **18**, 153–163.
- Bertelson, P. (1967). The time course of preparation. *Quarterly Journal of Experimental Psychology*, **19**, 272–279.
- Borger, R. (1963). The refractory period and serial choice-reactions. *Quarterly Journal of Experimental Psychology*, **15**, 1–12.
- Broadbent, D. E. (1971). *Decision and stress*. London: Academic Press.
- Carrier, L. M., & Pashler, H. (1995). Attentional limits in memory retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **21**, 1339–1348.
- Craik, K. J. 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. J. 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.
- Creamer, L. R. (1963). Event uncertainly, psychological refractory period and human data processing. *Journal of Experimental Psychology*, **66**, 187–194.
- Davis, R. (1957). The human operator as a single channel information system. *Quarterly Journal of Experimental Psychology*, **9**, 119–129.
- Davis, R. (1959). The role of “attention” in the psychological refractory period. *Quarterly Journal of Experimental Psychology*, **11**, 211–220.
- De Jong, R. (1993). Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 965–980.
- Duncan, J. (1980). The demonstration of capacity limitation. *Cognitive Psychology*, **12**, 75–96.
- Ehrenstein, A., Schweikert, R., Choi, S., & Proctor, R. W. (1997). Scheduling processes in

- working memory: Instructions control the order of memory search and mental arithmetic. *The Quarterly Journal of Experimental Psychology*, **50A**, 766–802.
- 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.
- Gopher, D. (1986). In defence of resources: On structures, energies, pools and the allocation of attention. In G. R. J. Hockey, A. W. K. Gaillard, & M. G. H. Coles (Eds.) *Energetics and human information processing* (pp. 353–371). Dordrecht: Martinus Nijhoff.
- Gopher, D. (1993). The skill of attention control: Acquisition and execution of attention strategies. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology artificial intelligence, and cognitive neuroscience* (pp. 299–322). Cambridge, MA: MIT Press.
- Gopher, D. (1994). Analysis and measurement of mental load. In G. d'Ydewalle, P. Eelen, & P. Bertelson (Eds.), *International perspectives on psychological science. Vol 2: The state of the art* (pp. 265–291). Hove, UK: Erlbaum.
- Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (pp. 1–49). New York: Wiley.
- Gopher, D., & Navon, D. (1980). How is performance limited: Testing the notion of central capacity. *Acta Psychologica*, **46**, 161–180.
- Gopher, D., Brickner, M., & Navon, D. (1982). Different difficulty manipulations interact differently with task emphasis: Evidence for multiple resources, **8**, 146–157.
- Gottsdanker, R. (1979). A psychological refractory period or an unprepared period? *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 208–215.
- Gottsdanker, R. (1980). The ubiquitous role of preparation. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 355–371). Amsterdam: North Holland.
- Gottsdanker, R., & Way, T. C. (1966). Varied and constant intersignal intervals in psychological refractoriness. *Journal of Experimental Psychology*, **72**, 792–804.
- Hawkins, H. L., Rodriguez, E., & Reicher, G. M. (1979). *Is time-sharing a general ability?* ONR Technical Report No. 3, University of Oregon, Eugene, OR.
- Herman, L. M., & Kantowitz, B. H. (1970). The psychological refractory period effect: Only half the double stimulation story? *Psychological Bulletin*, **73**, 74–88.
- Hirst, W., & Kalmar, D. (1987). Characterizing attentional resources. *Journal of Experimental Psychology: General*, **116**, 68–81.
- Hirst, W., Spelke, E. S., Reaves, C. C., Caharack, G., & Neisser, U. (1980). Dividing attention without alternation or automaticity. *Journal of Experimental Psychology: General*, **109**, 98–117.
- Hommel, B. (1998). Automatic stimulus-response translation in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 1368–1384.
- James, W. (1890). *The principles of psychology*. New York: Holt.
- Johnston, J. C., McCann, R. S., & Remington, R. W. (1995). Chronometric evidence for two types of attention. *Psychological Science*, **6**, 365–369.
- Kafry, D. (1971). *Task complexity and refractoriness in the RSI design*. Unpublished M.A. thesis, Hebrew University, Jerusalem (in Hebrew, cited in Kahneman, 1973).
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Kantowitz, B. H. (1974). Double stimulation with varying response requirements. *Journal of Experimental Psychology*, **103**, 1092–1107.

- Kantowitz, B. H., & Knight, J. L. (1976). Testing tapping time-sharing. II. Auditory secondary task. *Acta Psychologica*, **40**, 343–362.
- Kantowitz, B. H., & Knight, J. L. (1976). On experimenter-limited processes. *Psychological Review*, **83**, 502–507.
- 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*. Palisades, CA: Goodyear.
- Kerr, B. (1973). Processing demands during mental operations. *Memory & Cognition*, **1**, 401–412.
- Kinchla, R. A. (1980). The measurement of attention. In R. Nickerson (Ed.), *Attention and Performance VIII* (pp. 213–238). Hillsdale, NJ: Erlbaum.
- Kinsbourne, M. (1981). Single channel theory. In D. H. Holding (Ed.), *Human skills* (pp. 65–89). Chichester, UK: Wiley.
- Koch, R. (1995). Hicks's law and the psychological refractory period. Paper presented at the KNAW Symposium on Discrete versus Continuous Information Processing, Amsterdam, The Netherlands.
- Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance*, **5**, 189–207.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, **95**, 492–527.
- Logan, G. D., & Burkell, J. (1986). Dependence and independence in responding to double stimulation: A comparison of stop, change, and dual-task paradigms. *Journal of Experimental Psychology: Human Perception and Performance*, **12**, 549–563.
- Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, **108**, 393–434.
- Logan, G. D., & Schulkind, M. D. (2000). Parallel memory retrieval in dual-task situations. I. Semantic memory. *Journal of Experimental Psychology: Human Perception and Performance*, **26**, 1072–1090.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. Oxford, UK: Oxford Univ. Press.
- 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.
- McGill, W. (1963). Stochastic latency mechanisms. In R. D. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 309–360). New York: Wiley.
- McLeod, P. (1977). Parallel processing and the psychological refractory period. *Acta Psychologica*, **41**, 381–391.
- Meyer, D. E., & Kieras, D. E. (1997a). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, **104**, 3–65.
- Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory-period phenomena. *Psychological Review*, **104**, 749–791.
- Meyer, D. E., & Kieras, D. E. (1999). Precise to a practical theory of cognition and action: Some lessons from EPIC computational models of human multiple-task performance. In D. Gopher & A. Koriati (Eds.), *Attention and Performance XVII* (pp. 17–88). Cambridge, MA: MIT Press.

- Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica*, **27**, 84–92.
- Navon, D. (1984). Resources—A theoretical soup stone? *Psychological Review*, **91**, 216–234.
- Navon, D. (1985). Attention division or attention sharing? In M. I. Posner & O. S. M. Marin (Eds.), *Attention and Performance XI* (pp. 133–146). Hillsdale, NJ: Erlbaum.
- Navon, D. (1989a). The importance of being visible: On the role of attention in a mind viewed as an anarchic intelligence system; Basic tenets. *European Journal of Cognitive Psychology*, **1**, 191–213.
- Navon, D. (1989b). The importance of being visible: On the role of attention in a mind viewed as an anarchic intelligence system; Applications to the field of attention. *European Journal of Cognitive Psychology*, **1**, 215–238.
- Navon, D. (1990a). Do people allocate limited processing resources among concurrent activities? In L. Green & J. H. Kagel (Eds.), *Advances in behavioral economics* (Vol. 2, pp. 209–225). Norwood, NJ: Ablex.
- Navon, D. (1990b). Exploring two methods for estimating performance trade-off. *Bulletin of the Psychonomic Society*, **28**, 155–157.
- Navon, D. (1998). Is it processing capacity that is being defined? *Behavioral and Brain Sciences*, **21**, 841–842.
- Navon, D., & Gopher, D. (1979). On the economy of the human information processing system. *Psychological Review*, **86**, 214–255.
- Navon, D., & Gopher, D. (1980). Task difficulty, resources, and dual-task performance. In R. S. Nickerson (Ed.), *Attention and Performance* (Vol. 8, pp. 297–315). Hillsdale, NJ: Erlbaum.
- Navon, D., Gopher, D., Chillag, N., & Spitz, G. (1984). On separability of and interference between tracking dimensions in dual-axis tracking. *Journal of Motor Behavior*, **16**, 364–391.
- Navon, D., & Miller, J. (1987). The role of outcome conflict in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, **13**, 435–448.
- Navon, D., & Miller, J. (1998). *How single is the single bottleneck?* Unpublished manuscript, Dunedin, New Zealand, University of Otago.
- Norman, D. A., & Bobrow, D. J. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, **7**, 44–64.
- Pashler, H. E. (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 358–377.
- Pashler, H. E. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two component theory of divided attention in simple tasks. *Cognitive Psychology*, **21**, 469–514.
- Pashler, H. E. (1990). Do response modality effects support multiprocessor models of divided attention? *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 826–842.
- Pashler, H. E. (1991). Shifting visual attention and selecting motor responses: Distinct attentional mechanisms. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 1023–1040.
- Pashler, H. E. (1993). Doing two things at the same time. *American Scientist*, **81**, 48–55.
- 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: Human Experimental Psychology*, **47**, 161–191.
- Pashler, H. E. (1998). *The psychology of attention*. Cambridge, MA: The MIT Press.
- Pashler, H. E., & 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.
- Pashler, H. E., & O'Brien, S. (1993). Dual-task interference and the cerebral hemispheres. *Journal of Experimental Psychology: Human Perception and Performance*, **19**, 315–330.
- Popper, K. R. (1995). *The logic of scientific discovery* (14th ed.). London: Routledge. [First published in 1934.]
- Posner, M. I. (1978). *Chronometric explorations of the mind*. Hillsdale, NJ: Erlbaum.
- Posner, M. I., & Boies, S. W. (1971). Components of attention. *Psychological Review*, **78**, 391–408.
- Rabbitt, P. M. A. (1969). Psychological refractory delay and response–stimulus interval duration in serial, choice-response tasks. *Acta Psychologica*, **30**, 195–219.
- Roldan, C. E. (1979). *Time-shared comparatory tracking and attention*. Paper presented at the 40th meeting of the Canadian Psychological Association, Quebec City.
- Rosenbrock, H. H. (1960). An automatic method for finding the greatest or least value of a function. *Computer Journal*, **3**, 175–184.
- Ruthruff, E. R., Pashler, H. E., & Klaassen, A. (1996). Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement. Unpublished manuscript.
- Sanders, A. F. (1979). Some remarks on mental load. In N. Moray (Ed.), *Mental workload: Its theory and measurement* (pp. 41–77). New York: Plenum Press.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing. I. Detection, search and attention. *Psychological Review*, **84**, 1–66.
- Shumacher, E. H., Lauber, E. J., Glass, J. M., Zurbruggen, E. L., Gmeindl, L., Kieras, D. E., & Meyer, D. E. (1999). Concurrent response-selection processes in dual-task performance: Evidence for adaptive executive control of task scheduling. *Journal of Experimental Psychology: Human Perception and Performance*, **25**, 791–814.
- Smith, M. C. (1967). The psychological refractory period as a function of performance of a first response. *Quarterly Journal of Experimental Psychology*, **19**, 350–352.
- Smith, M. C. (1969). The effect of varying information on the psychological refractory period. *Acta Psychologica*, **30**, 220–231.
- Sperling, G., & Melchner, M. J. (1978). The attention operation characteristics: Examples from visual search. *Science*, **202**, 315–318.
- Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology*, **14**, 1–36.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge Univ. Press.
- Triggs, T. J. (1968). *Capacity sharing and speeded reactions to successive signals*. Doctoral dissertation, University of Michigan, 1968. Dissertation Abstracts International 1969 30 877B (University Microfilms No. 69-12, 259).
- Tsang, P. S., Shaner, T. L., & Vidulich, M. A. (1995). Resource scarcity and outcome conflict in time-sharing performance. *Perception and Psychophysics*, **57**, 365–378.
- Ulrich, R., & Stapf, K. H. (1985). A double-response paradigm to study stimulus intensity

- effects upon the motor system in simple reaction time experiments. *Perception and Psychophysics*, **36**, 545–558.
- van der Heijden, A. H. C. (1981). *Short-term visual information forgetting*. London: Routledge & Kegan Paul.
- van Selst, M., Ruthruff, E., & Johnston, J. (1999). Can practice eliminate the psychological refractory period effect? *Journal of Experimental Psychology: Human Perception and Performance*, **25**, 1268–1283.
- Welford, A. T. (1952). The “psychological refractory period” and the timing of high-speed performance—A review and a theory. *British Journal of Psychology*, **43**, 2–19.
- Welford, A. T. (1967). Single-channel operation in the brain. *Acta Psychologica*, **27**, 5–22.
- Wickens, C. D. (1980). The structure of attentional resources. In R. S. Nickerson (Ed.), *Attention and performance* (Vol. XIII, pp. 239–257). Hillsdale, NJ: Erlbaum.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–102). New York: Academic Press.
- (Accepted May 14, 2001; published online November 7, 2001)